LETTER

Genomic epidemiology reveals multiple introductions of Zika virus into the United States

Nathan D. Grubaugh[1]*, Jason T. Ladner[2]*, Moritz U. G. Kraemer[3,4,5]*, Gytis Dudas[6]*, Amanda L. Tan[7]*, Karthik Gangavarapu[1]*, Michael R. Wiley[2,8]*, Stephen White[9]*, Julien Thézé[3]*, Diogo M. Magnani[10], Karla Prieto[2,8], Daniel Reyes[2,11], Andrea M. Bingham[12], Lauren M. Paul[7], Refugio Robles-Sikisaka[1], Glenn Oliveira[13], Darryl Pronty[9], Carolyn M. Barcellona[7], Hayden C. Metsky[14], Mary Lynn Baniecki[14], Kayla G. Barnes[14], Bridget Chak[14], Catherine A. Freije[14], Adrianne Gladden-Young[14], Andreas Gnirke[14], Cynthia Luo[14], Bronwyn MacInnis[14], Christian B. Matranga[14], Daniel J. Park[14], James Qu[14], Stephen F. Schaffner[14], Christopher Tomkins-Tinch[14], Kendra L. West[14], Sarah M. Winnicki[14], Shirlee Wohl[14], Nathan L. Yozwiak[14], Joshua Quick[15], Joseph R. Fauver[16], Kamran Khan[17,18], Shannon E. Brent[17], Robert C. Reiner Jr[19], Paola N. Lichtenberger[20], Michael J. Ricciardi[10], Varian K. Bailey[10], David I. Watkins[10], Marshall R. Cone[21], Edgar W. Kopp IV[21], Kelly N. Hogan[21], Andrew C. Cannons[21], Reynald Jean[22], Andrew J. Monaghan[23], Robert F. Garry[24], Nicholas J. Loman[15], Nuno R. Faria[3], Mario C. Porcelli[25], Chalmers Vasquez[25], Elyse R. Nagle[2,11], Derek A. T. Cummings[26], Danielle Stanek[12], Andrew Rambaut[27,28], Mariano Sanchez-Lockhart[2,11], Pardis C. Sabeti[14,29,30,31]§, Leah D. Gillis[9]§, Scott F. Michael[7]§, Trevor Bedford[6]§, Oliver G. Pybus[3]§, Sharon Isern[7]§, Gustavo Palacios[2]§ & Kristian G. Andersen[1,13,32]§

Zika virus (ZIKV) is causing an unprecedented epidemic linked to severe congenital abnormalities[1,2]. In July 2016, mosquito-borne ZIKV transmission was reported in the continental United States; since then, hundreds of locally acquired infections have been reported in Florida[3,4]. To gain insights into the timing, source, and likely route(s) of ZIKV introduction, we tracked the virus from its first detection in Florida by sequencing ZIKV genomes from infected patients and *Aedes aegypti* mosquitoes. We show that at least 4 introductions, but potentially as many as 40, contributed to the outbreak in Florida and that local transmission is likely to have started in the spring of 2016—several months before its initial detection. By analysing surveillance and genetic data, we show that ZIKV moved among transmission zones in Miami. Our analyses show that most introductions were linked to the Caribbean, a finding corroborated by the high incidence rates and traffic volumes from the region into the Miami area. Our study provides an understanding of how ZIKV initiates transmission in new regions.

ZIKV transmission in the Americas was first reported in Brazil in May 2015[5], although the virus was probably introduced 1–2 years earlier[6–8]. By January 2016, ZIKV cases had been reported in several South and Central American countries and most islands in the Caribbean[9]. Like dengue virus (DENV) and chikungunya virus (CHIKV), ZIKV is vectored primarily by *Aedes* mosquitoes[10–13]. The establishment of the peridomestic species *A. aegypti* in the Americas[14]

has facilitated the establishment of DENV, CHIKV, and now probably ZIKV as endemic in this region[15]. In the continental United States, transient outbreaks of DENV and CHIKV have been reported in regions of Texas and Florida[4,16–21] with abundant seasonal *A. aegypti* populations[14,22].

The 2016 ZIKV outbreak in Florida generated 256 confirmed ZIKV infections[4] (Fig. 1a). While transmission was confirmed across four counties in Florida (Fig. 1b), the outbreak was most intense in Miami-Dade County (241 infections). Although the case location could not always be determined, at least 114 (47%) infections are likely to have been acquired in one of three distinct transmission zones: Wynwood, Miami Beach, and Little River (Fig. 1c, d).

Using mosquito surveillance data, we determined the extent of mosquito-borne ZIKV transmission in Miami. Of the 24,351 mosquitoes collected from June to November 2016, 99.8% were *A. aegypti* and 8 pools of ≤50 mosquitoes tested positive for ZIKV (Fig. 1c, Extended Data Fig. 1). From these pools, we estimated that approximately 1 out of 1,600 *A. aegypti* mosquitoes were infected (0.061%, 95% confidence interval (CI) 0.028–0.115%, Extended Data Fig. 1a). This is similar to infection rates during DENV and CHIKV outbreaks[23]. Although we did not detect ZIKV-infected mosquitoes outside Miami Beach (Fig. 1c), we found that the number of human ZIKV cases correlated strongly with *A. aegypti* abundance within each transmission zone (Spearman $r = 0.61$, Fig. 1d, Extended Data Fig. 1b). This suggests that

[1]Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, California 92037, USA. [2]Center for Genome Sciences, US Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland 21702, USA. [3]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. [4]Boston Children's Hospital, Boston, Massachusetts 02115, USA. [5]Harvard Medical School, Boston, Massachusetts 02115, USA. [6]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. [7]Department of Biological Sciences, College of Arts and Sciences, Florida Gulf Coast University, Fort Myers, Florida 33965, USA. [8]College of Public Health, University of Nebraska Medical Center, Omaha, Nebraska 68198, USA. [9]Bureau of Public Health Laboratories, Division of Disease Control and Health Protection, Florida Department of Health, Miami, Florida 33125, USA. [10]Department of Pathology, University of Miami Miller School of Medicine, Miami, Florida 33136, USA. [11]Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, Nebraska 68198, USA. [12]Bureau of Epidemiology, Division of Disease Control and Health Protection, Florida Department of Health, Tallahassee, Florida 32399, USA. [13]Scripps Translational Science Institute, La Jolla, California 92037, USA. [14]The Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [15]Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, UK. [16]Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, Colorado 80523, USA. [17]Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Ontario M5B 1T8, Canada. [18]Division of Infectious Diseases, Department of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [19]Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington 98121, USA. [20]Division of Infectious Diseases, University of Miami Miller School of Medicine, Miami, Florida 33136, USA. [21]Bureau of Public Health Laboratories, Division of Disease Control and Health Protection, Florida Department of Health, Tampa, Florida 33612, USA. [22]Florida Department of Health in Miami-Dade County, Miami, Florida 33125, USA. [23]National Center for Atmospheric Research, Boulder, Colorado 80307, USA. [24]Department of Microbiology and Immunology, Tulane University School of Medicine, New Orleans, Louisiana 70112, USA. [25]Miami-Dade County Mosquito Control, Miami, Florida 33178, USA. [26]Department of Biology and Emerging Pathogens Institute, University of Florida, Gainesville, Florida 32610, USA. [27]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3FL, UK. [28]Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, USA. [29]Center for Systems Biology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [30]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts 02115, USA. [31]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. [32]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, USA.
*These authors contributed equally to this work.
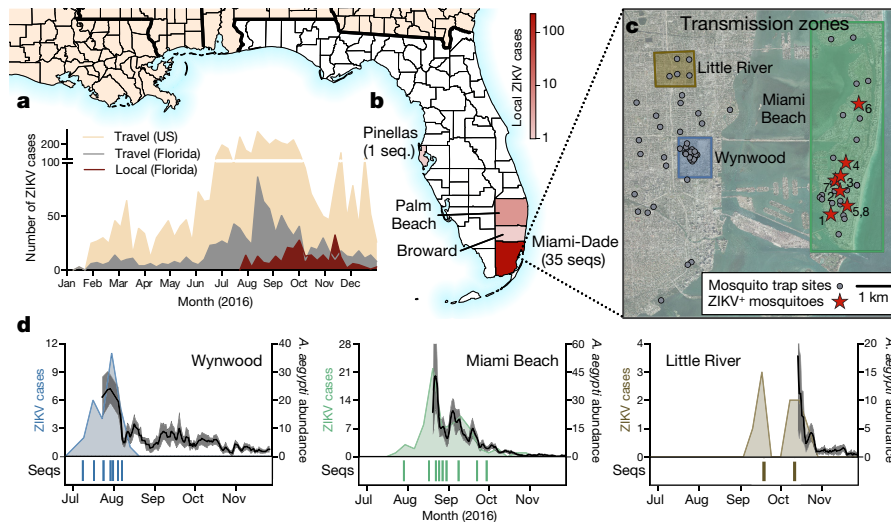§These authors jointly supervised this work.

Figure 1 | Zika virus outbreak in Florida. a, Weekly counts of confirmed travel-associated and locally acquired ZIKV cases in 2016. **b**, Four counties reported locally acquired ZIKV cases in 2016: Miami-Dade (241), Broward (5), Palm Beach (8), and Pinellas (1). There was also one case of unknown origin. **c**, The locations of mosquito traps and collected *A. aegypti* mosquitoes found to contain ZIKV RNA (ZIKV⁺) in relation to the transmission zones within Miami. **d**, Temporal distribution of weekly ZIKV cases (left *y*-axis), sequenced cases (bottom), and *A. aegypti* abundance per trap night (right *y*-axis) associated with the three described transmission zones. ZIKV cases and sequences are plotted in relation to symptom onset dates ($n = 18$). Sequenced cases without onset dates or that occurred outside the transmission zones are not shown ($n = 10$). Human cases and *A. aegypti* abundance per week were positively correlated (Spearman $r = 0.61$, Extended Data Fig. 1b). The maps were generated using open source basemaps (http://www.esri.com/data/basemaps).

A. aegypti mosquitoes were the primary mode of transmission and that changes to vector abundance affected human infection rates. We found that the application of insecticides[3] suppressed mosquito populations during periods of intensive use (Extended Data Fig. 1c), and therefore probably contributed to ZIKV clearance.

We sequenced 39 ZIKV genomes from clinical and mosquito samples without cell culture[24] (Supplementary Table 1a). Our ZIKV dataset included 29 genomes from patients with locally acquired infections (Fig. 1d) and 7 from *A. aegypti* pools (Fig. 1c). We also sequenced three ZIKV genomes from travel-associated cases in Florida. Our dataset included cases from all transmission zones in Miami (Fig. 1d) and represented about 11% of all confirmed locally acquired cases in Florida. We made all sequence data openly available in the NCBI BioProject database (PRJNA342539, PRJNA356429) immediately after data generation.

We reconstructed phylogenetic trees from our ZIKV genomes along with 65 published genomes from other affected regions (Fig. 2, Extended Data Figs 2, 3). We found that the Florida ZIKV genomes formed four distinct lineages (labelled F1–F4 in Fig. 2a), three of which (F1–F3) belonged to the same clade (labelled A in Fig. 2a). We sampled only a single human case each from the F3 and F4 lineages, consistent with limited transmission (Fig. 2a). The other two Florida lineages (F1 and F2) comprised ZIKV genomes from human and mosquito samples within Miami-Dade County (Fig. 2b).

Using time-structured phylogenies[25], we estimated that at least four separate introductions were responsible for the locally acquired cases observed in our dataset. The phylogenetic placement of lineage F4 clearly indicates that it resulted from an independent introduction of a lineage distinct from those in clade A (Fig. 2a). For the two well-supported nodes linking lineages F1 and F2 (labelled B, Fig. 2a) and F1–F3 (A, Fig. 2a), we estimated the time of the most recent common ancestor (tMRCA) to be during the summer of 2015 (95% highest posterior density (HPD) June–September 2015). Our data displayed a strong clock signal (Extended Data Fig. 2b) and tMRCA estimates were robust across a range of models (Extended Data Table 1a). Thus, although lineages F1–F3 belong to clade A, any fewer than three distinct introductions leading to these lineages would have required undetected transmission of ZIKV in Florida for approximately one year (Fig. 2a).

To estimate the likelihood of a single ZIKV transmission chain persisting for more than a year, we modelled spread under different assumptions of the basic reproductive number (R_0). Using the number of locally acquired and travel-associated cases, along with the number of observed genetic lineages, we estimated an R_0 between 0.5 and 0.8 in Miami-Dade County (Extended Data Fig. 4). Even at the upper end of this range, the probability of a single transmission chain persisting for over a year is extremely low (~0.5%, Fig. 2c). This is especially true considering the low *A. aegypti* abundance during the winter months (Extended Data Fig. 1d).

Given the low probability of long-term persistence, we expect that our ZIKV genomes (F1–F4) were the result of at least four introductions. Differences in surveillance practices and a high number of travel-associated cases (Fig. 1a), however, probably mean that unsampled ZIKV introductions also contributed to the outbreak. To estimate the total number of ZIKV introductions, we modelled scenarios that resulted in 241 locally acquired cases within Miami-Dade County and found that, with R_0 values of 0.5–0.8, we expect 17–42 (95% CI 3–63) separate introductions to have contributed to the outbreak (Fig. 2d). The majority of these introductions would be likely to have generated a single secondary case that was undetected in our genetic sampling (Extended Data Fig. 4a). Incorporating under-reporting in a sensitivity analysis increases R_0 estimates slightly to 0.7–0.9 (Extended Data Fig. 4f–i).

The two main ZIKV lineages, F1 and F2, included the majority of genomes from Florida (92%, Fig. 2a). Assuming that they represent two independent introductions, we estimated when each of these lineages arrived in Florida. The probability densities for the tMRCAs of both F1 and F2 were centred around March–April 2016 (Fig. 2b, 95% HPD January–May 2016). The estimated timing for these introductions corresponds with the presence of suitable *A. aegypti* populations in Miami-Dade County[26] (Extended Data Fig. 1d) and suggests that ZIKV transmission could have started at least 2 months before its detection in July 2016 (Fig. 1a). The dates of the introductions could be more recent if multiple F1 or F2 lineage viruses arrived independently. However, more than two introductions would be necessary to substantially change our estimates for the timing of the earliest introduction.
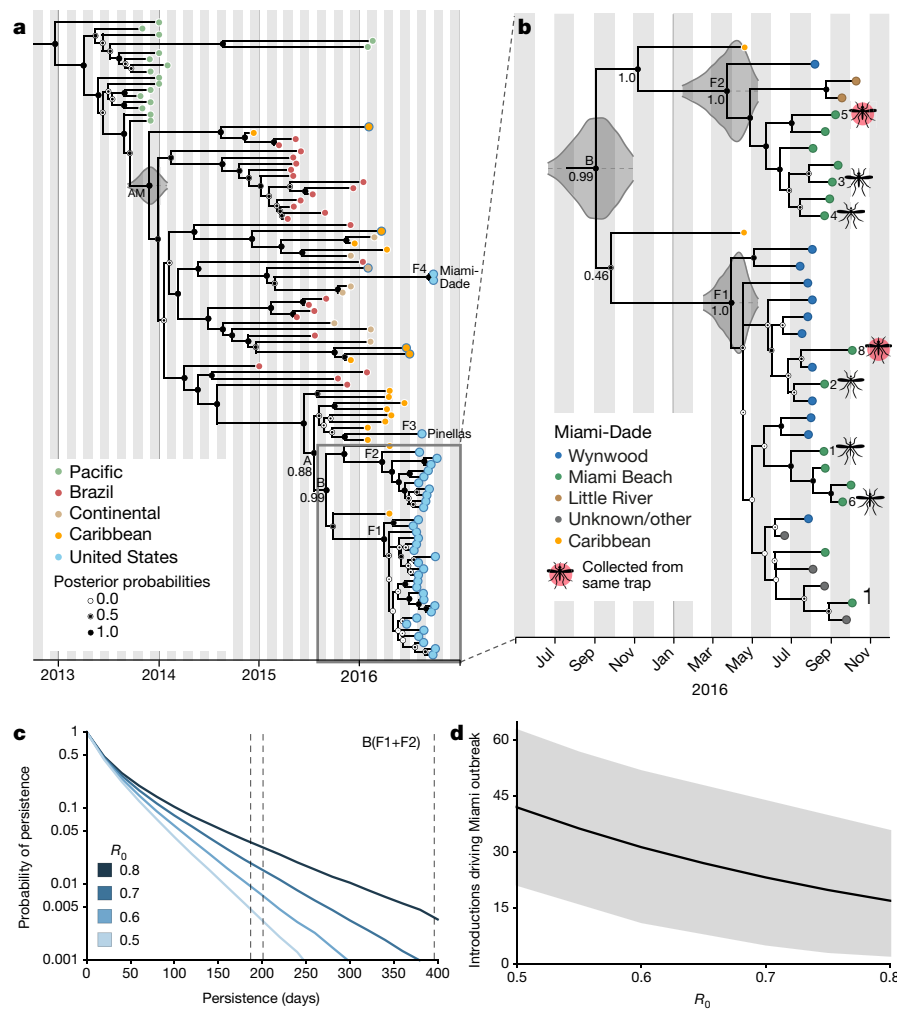
Figure 2 | Multiple introductions of Zika virus into Florida.
a, Maximum clade credibility (MCC) tree of ZIKV genomes sequenced from outbreaks in the Pacific islands and the epidemic in the Americas. Tips are coloured according to collection location. The five tips outlined in blue but filled with a different colour indicate ZIKV cases in the US associated with travel (fill colour indicates the probable location of infection). Clade posterior probabilities are indicated by white circles filled with black relative to the level of support. The grey violin plot indicates the 95% HPD interval for the tMRCA for the epidemic in the Americas (AM). Lineage F4 contains two identical ZIKV genomes from the same patient.

b, A zoomed-in version of the whole MCC tree showing the collection locations of Miami-Dade sequences and whether they were sequenced from mosquitoes (numbers correspond to trap locations in Fig. 1c). 95% HPD intervals are shown for the tMRCAs. **c**, The probability of ZIKV persistence after introduction for different R_0 values. Persistence is measured as the number of days from initial introduction of viral lineages until their extinction. Vertical dashed lines show the inferred mean persistence time for lineages F1, F2 and B based on their tMRCAs. **d**, Total number of introductions (mean with 95% CI) that contributed to the outbreak of 241 local cases in Miami-Dade County for different R_0 values.

To understand transmission dynamics within Miami, we analysed our genomic data together with case data from the Florida Department of Health (Supplementary Table 1a). Although the three ZIKV transmission zones were spatially distinct, they occurred within about 5 km of each other (Fig. 1c) and the ZIKV infections associated with each zone overlapped temporally (Fig. 1d). Our ZIKV genomes with zone assignments all belonged to lineages F1 and F2, but neither of these lineages was confined to a single zone (Fig. 2b). In fact, we detected both F1 and F2 lineage viruses from *A. aegypti* collected from the same trap 26 days apart (mosquitoes 5 and 8, Fig. 2b). These findings suggest that ZIKV moved among areas of Miami.

Determining the sources and routes of ZIKV introductions could help to mitigate future outbreaks. We found that lineages F1–F3 clustered with ZIKV genomes sequenced from the Dominican Republic and Guadeloupe (Fig. 2, Extended Data Figs 2, 3). By contrast, F4 clustered with genomes from Central America (Fig. 2, Extended Data Figs 2, 3). These findings suggest that whereas ZIKV outbreaks occurred throughout the Americas, the Caribbean islands were the main source of local ZIKV transmission in Florida. Because of severe undersampling

of ZIKV genomes, however, we cannot rule out other source areas. Similarly, even though we found that the Florida ZIKV genomes clustered together with sequences from the Dominican Republic, our results do not prove that ZIKV entered Florida from this country.

We investigated ZIKV infection rates and travel patterns to corroborate our phylogenetic evidence for Caribbean introductions. We found that the Caribbean islands bore the highest ZIKV incidence rates (Fig. 2b), despite Brazil and Colombia reporting the highest absolute number of cases (January–June 2016; Fig. 3a, Extended Data Fig. 5, Supplementary Table 1b). During the same time period, we estimated that about 3 million travellers arrived from the Caribbean, accounting for 54% of the total traffic into Miami, with the vast majority (about 2.4 million) arriving via cruise ships (Fig. 3b, Extended Data Fig. 6, Supplementary Table 1b). Combining the infection rates with travel capacities, we estimated that around 60–70% of ZIKV-infected travellers arrived from the Caribbean (Fig. 3c, Extended Data Fig. 7a). We also found that the number of travel-associated ZIKV cases correlated strongly with the expected number of importations from the Caribbean (Spearman $r = 0.8$; Fig. 3d, Extended Data Fig. 7b). Finally,
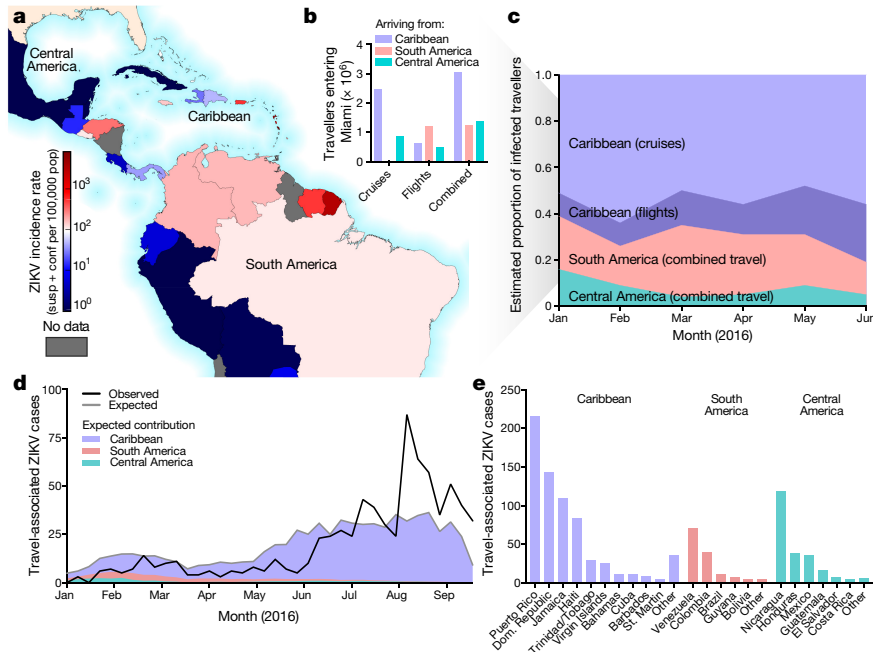
Figure 3 | Frequent opportunities for Zika virus introductions into Miami from the Caribbean. a, Reported ZIKV cases per country or territory from January to June 2016, normalized by total population. **b**, The number of estimated travellers entering Miami during January–June 2016, by method of travel. **c**, The number of travellers and reported ZIKV incidence rate for the country or territory of origin were used to estimate the proportion of infected travellers coming from each region with ZIKV in the Americas. **d**, The observed number of weekly travel-associated ZIKV cases in Florida (black line), plotted with the expected number of ZIKV-infected travellers (as estimated in **c**) coming from all of the Americas (grey line) and the regional contributions (coloured areas). **e**, The countries visited by the 1,016 patients with travel-associated ZIKV diagnosed in Florida.

67% of individuals with travel-associated infections in Florida reported recent travel to the Caribbean (Fig. 3e); however, their mode of travel is unknown. Together, these findings suggest that a high incidence of ZIKV in the Caribbean, combined with frequent travel, could have played a key role in the establishment of ZIKV transmission in Florida. These findings, however, do not indicate that cruise ships themselves are risk factors for human ZIKV infection, but only that they served as a major mode of transportation from areas with active transmission. In addition, ZIKV exposure may vary among individuals depending on their purpose of travel and therefore we cannot determine the specific contribution of ZIKV-infected travellers arriving via airlines or cruise ships.

The majority of the Florida ZIKV outbreak occurred in Miami-Dade County (Fig. 1b). To determine whether there is a higher potential for ZIKV outbreaks in this area, we analysed incoming passenger traffic from regions with ZIKV transmission along with local *A. aegypti* abundance. We estimated that Miami and nearby Fort Lauderdale received

around 72% of traffic (Fig. 4) and that Miami received more air and sea traffic from ZIKV-endemic areas than any other city in the United States (Extended Data Fig. 8). We estimated that, during January–April 2016, *A. aegypti* abundance was highest in southern Florida[22] (Fig. 4, Extended Data Figs 1d, 8). By June, most of Florida and several cities across the South probably supported high populations of *A. aegypti*[14,22] (Extended Data Fig. 8); however, most of this region has not reported local *A. aegypti*-borne virus transmission for at least 60 years[19]. In fact, the only region outside Florida with local ZIKV transmission is southern Texas[27], which is also the only other region with recent DENV outbreaks[19–21]. Therefore, the combination of travellers, mosquito ecology, and human population density is likely to make Miami one of the few places in the continental United States at risk for *A. aegypti*-borne virus outbreaks[22,26,28].

The extent of ZIKV transmission in Florida was unprecedented, with more reported ZIKV cases in 2016 (256) than DENV cases since 2009 (136)[4,16,17]. This case difference may be reflected by lower incidence of
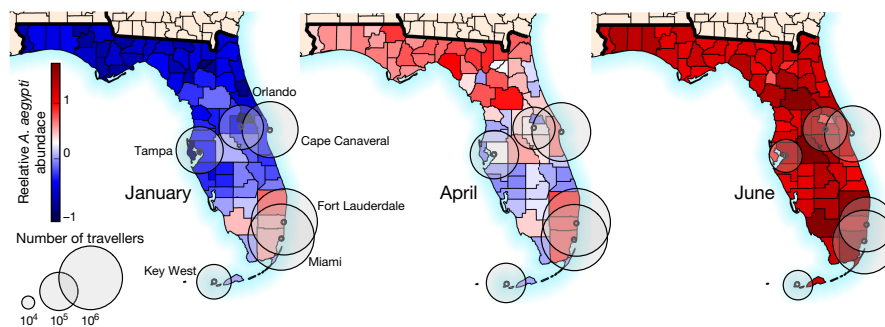

Figure 4 | Southern Florida has a high potential for *A. aegypti*-borne virus outbreaks. The estimated number of travellers per month (circles) entering Florida cities via flights and cruise ships, plotted with estimated relative *A. aegypti* abundance. Only cities receiving more than 10,000 passengers per month are shown. Relative *A. aegypti* abundance for every month is shown in Extended Data Fig. 1d.

endemic DENV than epidemic ZIKV in source countries[29,30], resulting in fewer DENV importations (reported travel-associated cases since 2009: 654 DENV and 1,016 ZIKV)[4]. Given that the majority of ZIKV infections are asymptomatic[2,31], the true number of ZIKV cases is likely to have been much higher. Despite this, we estimated that the average R_0 was less than 1 and therefore multiple introductions were necessary to give rise to the observed outbreak[32]. The high volume of traffic entering Florida from ZIKV-affected regions, especially the Caribbean, is likely to have provided a substantial supply of ZIKV-infected individuals[33]. Because Florida is unlikely to sustain long-term ZIKV transmission[32], the potential for future ZIKV outbreaks in this region is dependent upon activity elsewhere. Therefore, we expect that outbreaks in Florida will cycle with ZIKV transmission dynamics in the Americas[7,8,15].

1. Zika virus and complications. *World Health Organization* http://www.who.int/features/qa/zika/en/ (2016).
2. Lazear, H. M. & Diamond, M. S. Zika virus: new clinical syndromes and its emergence in the Western hemisphere. *J. Virol.* **90,** 4864–4875 (2016).
3. Likos, A. *et al.* Local mosquito-borne transmission of Zika virus—Miami-Dade and Broward Counties, Florida, June–August 2016. *MMWR Morb. Mortal. Wkly. Rep.* **65,** 1032–1038 (2016).
4. Mosquito-Borne Disease Surveillance. *Florida Department of Health* http://www.floridahealth.gov/diseases-and-conditions/mosquito-borne-diseases/surveillance.html (2017).
5. Hennessey, M., Fischer, M. & Staples, J. E. Zika virus spreads to new areas—region of the Americas, May 2015–January 2016. *MMWR Morb. Mortal. Wkly. Rep.* **65,** 55–58 (2016).
6. Faria, N. R. *et al.* Zika virus in the Americas: early epidemiological and genetic findings. *Science* **352,** 345–349 (2016).
7. Faria, N. R. *et al.* Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* http://dx.doi.org/10.1038/nature22401 (2017).
8. Metsky, H. C. *et al.* Zika virus evolution and spread in the Americas. *Nature* http://dx.doi.org/10.1038/nature22402 (2017).
9. Regional Zika Epidemiological Update (Americas). *Pan American Health Organization* http://www.paho.org/hq/index.php?option=com_content&view=article&id=11599&Itemid=41691&lang=en (2016).
10. Weger-Lucarelli, J. *et al.* Vector competence of American mosquitoes for three strains of Zika virus. *PLoS Negl. Trop. Dis.* **10,** e0005101 (2016).
11. Guerbois, M. *et al.* Outbreak of Zika virus infection, Chiapas State, Mexico, 2015, and first confirmed transmission by *Aedes aegypti* mosquitoes in the Americas. *J. Infect. Dis.* **214,** 1349–1356 (2016).
12. Ferreira-de-Brito, A. *et al.* First detection of natural infection of *Aedes aegypti* with Zika virus in Brazil and throughout South America. *Mem. Inst. Oswaldo Cruz* **111,** 655–658 (2016).
13. Chouin-Carneiro, T. *et al.* Differential susceptibilities of *Aedes aegypti* and *Aedes albopictus* from the Americas to Zika virus. *PLoS Negl. Trop. Dis.* **10,** e0004543 (2016).
14. Kraemer, M. U. G. *et al.* The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife* **4,** e08347 (2015).
15. Ferguson, N. M. *et al.* Countering the Zika epidemic in Latin America. *Science* **353,** 353–354 (2016).
16. Teets, F. D. *et al.* Origin of the dengue virus outbreak in Martin County, Florida, USA 2013. *Virol. Rep.* **1-2,** 2–8 (2014).
17. Graham, A. S. *et al.* Mosquito-associated dengue virus, Key West, Florida, USA, 2010. *Emerg. Infect. Dis.* **17,** 2074–2075 (2011).
18. Kendrick, K., Stanek, D. & Blackmore, C. Notes from the field: Transmission of chikungunya virus in the continental United States—Florida, 2014. *MMWR Morb. Mortal. Wkly. Rep.* **63,** 1137 (2014).
19. Bouri, N. *et al.* Return of epidemic dengue in the United States: implications for the public health practitioner. *Public Health Rep.* **127,** 259–266 (2012).
20. Ramos, M. M. *et al.* Epidemic dengue and dengue hemorrhagic fever at the Texas–Mexico border: results of a household-based seroepidemiologic survey, December 2005. *Am. J. Trop. Med. Hyg.* **78,** 364–369 (2008).
21. Murray, K. O. *et al.* Identification of dengue fever cases in Houston, Texas, with evidence of autochthonous transmission between 2003 and 2005. *Vector Borne Zoonotic Dis.* **13,** 835–845 (2013).
22. Monaghan, A. J. *et al.* On the seasonal occurrence and abundance of the Zika virus vector mosquito *Aedes aegypti* in the contiguous United States. *PLoS Curr.* http://dx.doi.org/10.1371/currents.outbreaks.50dfc7f46798675fc63e7d7da563da76 (2016).
23. Dzul-Manzanilla, F. *et al.* Evidence of vertical transmission and co-circulation of chikungunya and dengue viruses in field populations of *Aedes aegypti* (L.) from Guerrero, Mexico. *Trans. R. Soc. Trop. Med. Hyg.* **110,** 141–144 (2016).
24. Quick, J. *et al.* Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protocols* http://dx.doi.org/10.1038/nprot.2017.066 (2017).
25. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and BEAST 1.7. *Mol. Biol. Evol.* **29,** 1969–1973 (2012).
26. Robert, M. A. *et al.* Modeling mosquito-borne disease spread in U.S. urbanized areas: the case of dengue in Miami. *PLoS One* **11,** e0161365 (2016).
27. McCarthy, M. First US case of Zika virus infection is identified in Texas. *Br. Med. J.* **352,** i212 (2016).
28. Nelson, B. *et al.* Travel volume to the United States from countries and U.S. territories with local Zika virus transmission. *PLoS Curr.* http://dx.doi.org/10.1371/currents.outbreaks.ac6d0f8c9c35e88825c1a1147697531c (2016).
29. Dengue in Puerto Rico. *The Centers for Disease Control and Prevention* https://www.cdc.gov/dengue/about/inpuerto.html (2017).
30. Pan American Health Organization. Countries and territories with autochthonous transmission in the Americas reported in 2015–2017. http://www.paho.org/hq/index.php?option=com_content&view=article&id=11603&Itemid=41696&lang=en (2016).
31. Duffy, M. R. *et al.* Zika virus outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* **360,** 2536–2543 (2009).
32. Dinh, L., Chowell, G., Mizumoto, K. & Nishiura, H. Estimating the subcritical transmissibility of the Zika outbreak in the State of Florida, USA, 2016. *Theor. Biol. Med. Model.* **13,** 20 (2016).
33. Nunes, M. R. T. *et al.* Air travel is associated with intracontinental spread of dengue virus serotypes 1–3 in Brazil. *PLoS Negl. Trop. Dis.* **8,** e2769 (2014).

Author Contributions All contributions are listed in order of authorship. Designed the experiments: N.D.G., J.T.L., G.D., M.U.G.K., D.A.T.C., P.C.S., L.D.G., S.F.M., T.B., O.G.P., S.I., G.P., and K.G.A. Collected samples: A.L.T., S.W., D.M.M., A.M.B., L.M.P., D.P., C.M.B., P.N.L., M.J.R., V.K.B., D.I.W., M.R.C., E.W.K., K.N.H., A.C.C., R.J., M.C.P., C.V., D.S., L.D.G., S.F.M., and S.I. Performed the sequencing: N.D.G., M.R.W., K.P., D.R., R.R.-S., G.O., and E.R.N. Provided data, reagents, or protocols: N.D.G., J.T.L., G.D., M.U.G.K., K.G., M.R.W., R.R.-S., G.O., H.C.M., M.L.B., K.G.B., B.C., C.A.F., A.G.-Y., A.G., C.L., B.M., C.B.M., D.J.P., J. Q.U, S.F.S., C.T.-T., K.L.W., S.M.W., S.W., N.L.Y., J.Qui., J.R.F., K.K., S.E.B., A.J.M., R.F.G., N.J.L., M.C.P., C.V., P.C.S., S.F.M., and S.I. Analysed the data: N.D.G., J.T.L., G.D., M.U.G.K., K.G., J.T., J.R.F., R.C.R., N.R.F., D.A.T.C., A.R., M.S.-L., T.B., S.F.M, O.G.P., S.I., and K.G.A. Edited manuscript: G.D., M.U.G.K., J.T., S.F.S., A.R., T.B., O.G.P., S.I., and G.P. Wrote manuscript: N.D.G., J.T.L., and K.G.A. All authors read and approved the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to K.G.A. (andersen@scripps.edu) or G.P. (gustavo.f.palacios.ctr@mail.mil).

Reviewer Information *Nature* thanks K. St George, A. Wilder-Smith, M. Worobey and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Ethical statement. This work was evaluated and approved by the relevant Institutional Review Boards (IRBs) or Ethics Review Committees at The Scripps Research Institute (TSRI) and the US Army Medical Research Institute of Infectious Diseases (USAMRIID) Office of Human Use and Ethics. This work was conducted as part of the public health response in Florida and samples were collected under a waiver of consent granted by the Florida Department of Health (DOH) Human Research Protection Program. The work received a non-human subjects research designation (category 4 exemption) by the Florida DOH because this research was performed with leftover clinical diagnostic samples involving no more than minimal risk. All samples were de-identified before receipt by the study investigators.

Florida Zika virus case data. Weekly reports of international travel-associated and locally acquired ZIKV infections diagnosed in Florida were obtained from the Florida DOH mosquito-borne disease surveillance system[4]. Dates of symptom onset from the Miami transmission zones (Wynwood, Miami Beach, and Little River) determined by the Florida DOH investigation process were obtained from the ZIKV resource website[34] and daily updates[35]. International travel-associated ZIKV case counts in the United States (outside Florida) were obtained from the CDC[36]. The local and travel-associated ZIKV case numbers for Florida were obtained from the Florida DOH. The one local ZIKV infection diagnosed in Duval County was believed to have originated elsewhere in Florida. Therefore, this case is listed as 'unknown origin' in Fig. 1b. In Fig. 3e, only the countries visited five or more times by ZIKV-infected travellers diagnosed in Florida are shown. Countries with fewer than five visits were aggregated into an "other" category by region (that is, Caribbean, South America, or Central America).

Clinical sample collection and RNA extraction. Clinical samples from locally acquired ZIKV infections were collected between 22 June and 11 October 2016. The Florida DOH identified persons with compatible illness and clinical samples were shipped to the Bureau of Public Health Laboratories for confirmation by qRT–PCR and antibody tests following interim guidelines[3,37–39]. Clinical specimens (whole blood, serum, saliva, or urine) submitted for analysis were refrigerated or frozen at or below −70 °C until RNA was extracted. RNA was extracted using the RNAeasy kit (QIAGEN), MagMAX for Microarrays Total RNA Isolation Kit (Ambion), or MagNA Pure LC 2.0 or 96 Systems (Roche Diagnostics). Purified RNA was eluted into 50–100 μl using the supplied elution buffers, immediately frozen at or below −70 °C, and transported on dry ice. The Florida DOH also provided investigation data for these samples, including symptom onset dates and, when available, assignments to the zone where infection was likely to have occurred (Supplementary Table 1).

Mosquito collection, RNA extraction, and entomological data analysis. 24,351 *A. aegypti* and *Aedes albopictus* mosquitoes (sorted into 2,596 pools) were collected throughout Miami-Dade County during June–November 2016 using BG-Sentinel mosquito traps (Biogents AG). Up to 50 mosquitoes of the same species and sex were pooled per trap. The pooled mosquitoes were stored in RNAlater (Invitrogen), RNA was extracted using either the RNAeasy kit (QIAGEN) or MagMAX for Microarrays Total RNA Isolation Kit (Ambion), and ZIKV RNA was detected by qRT–PCR targeting the envelope protein coding region[39] or the Trioplex qRT–PCR kit[40]. ZIKV infection rates were calculated per 1,000 female *A. aegypti* mosquitoes using the bias-corrected maximum likelihood estimate (MLE)[41]. Days of insecticide usage by the Miami-Dade Mosquito Control were inferred from the zone-specific ZIKV activities timelines published by the Florida DOH[34].

Relative monthly *Aedes aegypti* abundance. For the purpose of this study we used *A. aegypti* suitability maps from ref. 14 and derived monthly estimates based on the statistical relationships between mosquito presence and environmental correlates[42]. Following ref. 43, we used a simple mathematical formula to transform the probability of detection maps into mosquito abundance maps. We assumed $P(Y=1)$ where Y is a binary variable (presence/absence). Using a Poisson distribution $X()$ to govern the abundance of mosquitoes, the probability of not observing any mosquitoes can be related to the probability of absence as: $P(X=0) = P(Y=0)$. We used the following transformation to generate abundance (λ) estimates per county in Florida:

$$e^{-\lambda} = P(Y=0)$$
$$\lambda = -\log(P(Y=0))$$
$$\lambda = -\log(1 - P(Y=1))$$

We did not consider *A. albopictus* abundance in this study because 99.8% of mosquitoes collected in Miami-Dade County were *A. aegypti*. Relative *A. aegypti*

abundance in major US cities presented in Extended Data Fig. 8 was estimated as previously described[22].

Zika virus quantification. ZIKV genome equivalents (GE) were quantified by qRT–PCR. At TSRI, ZIKV qRT–PCR was performed as follows: ZIKV RNA standards were transcribed from the ZIKV NS5 region (nucleotides (nt) 8,651–9,498) using the T7 forward primer (5′-TAATACGACTCACTATAGGGAGA TCAGGCTCCTGTCAAAACCC-3′), reverse primer (5′-AGTGACAACTTGTC CGCTCC-3′), and the T7 Megascript kit (Ambion). For qRT–PCR, primers and a probe targeting the NS5 region (nt 9,014–9,123) were designed using the ZIKV isolate PRVABC59 (GenBank: KU501215): forward primer (5′-AGTGCCAGAGC TGTGTGTAC-3′), reverse primer (5′-TCTAGCCCCTAGCCACATGT-3′), and FAM-fluorescent probe (5′-GGCAGCCGCGCCATCTGGT-3′). The qRT–PCR assays were performed in 25-μl reactions using the iScript One-step RT–PCR Kit for probes (Bio-Rad Laboratories Inc.) and 2 μl of sample RNA. Amplification was performed at 50 °C for 20 min, 95 °C for 3 min, and 40 cycles of 95 °C for 10 s and 57 °C for 10 s. Fluorescence was read at the end of the 57 °C annealing–extension step. Tenfold dilutions of the ZIKV RNA transcripts (2 μl per reaction) were used to create a standard curve for quantification of ZIKV GE per μl RNA. The lower limits of quantification are 4 GE per μl RNA, or at a cycle threshold of ~36.

ZIKV GE were quantified at USAMRIID using the University of Bonn ZIKV envelope protein (Bonn E) qRT–PCR assay[44]. RNA standards were transcribed using an amplicon generated from a ZIKV plasmid containing T7 promoter at the start of the 5′ untranslated region (UTR). The plasmid was designed using the ZIKV isolate BeH819015 (GenBank: KU365778.1) and the amplicon included nt 1–4,348, which covers the 5′ UTR, C, prM, M, E, NS1, and NS2 regions. The qRT–PCR assays were performed in 25-μl reactions using the SuperScript III platinum One-step qRT–PCR Kit (ThermoFisher) and 2 μl of sample RNA was used. Amplification was performed following conditions as previously described[44]. Tenfold dilutions of the ZIKV RNA transcripts (5 μl per reaction) were used to create a standard curve for quantification of ZIKV GE per μl RNA.

Amplicon-based Zika virus sequencing. ZIKV sequencing at TSRI was performed using an amplicon-based approach using the ZikaAsian V1 scheme, as described[24]. This approach is similar to 'RNA jackhammering' to sequence low-quality viral samples described in ref. 45. In brief, cDNA was reverse-transcribed from 5 μl RNA using SuperScript IV (Invitrogen). ZIKV cDNA (2.5 μl per reaction) was amplified in 35 × 400-bp fragments from two multiplexed PCR reactions using Q5 DNA High-fidelity Polymerase (New England Biolabs). The amplified ZIKV cDNA fragments (50 ng) were prepared for sequencing using the Kapa Hyper prep kit (Kapa Biosystems) and SureSelect XT2 indexes (Agilent). Agencourt AMPure XP beads (Beckman Coulter) were used for all purification steps. Paired-end 251-nt reads were generated on the MiSeq using the V2 500 cycle or V3 600 cycle kits (Illumina).

Trimmomatic was used to remove primer sequences (first 22 nt from the 5′ end of the reads, which is the maximum length of the primers used for the multiplexed PCR) and bases at both ends with Phred quality score <20 (ref. 46). The reads were then aligned to the complete genome of a ZIKV isolate from Dominican Republic, 2016 (GenBank: KU853012) using Novoalign v3.04.04 (www.novocraft.com). Samtools was used to sort the aligned BAM files and to generate alignment statistics[47]. Snakemake was used as the workflow management system[48]. The code and reference indexes for the pipeline can be found at https://github.com/andersen-lab/zika-pipeline. ZIKV-aligned reads were visually inspected using Geneious v9.1.5 (ref. 49) before generating consensus sequences. A minimum of 3 × read-depth coverage, in support of the consensus, was required to make a base call.

Enrichment-based Zika virus sequencing. ZIKV sequencing at USAMRIID was performed using a targeted enrichment approach. Sequencing libraries were prepared using the TruSeq RNA Access Library Prep kit (Illumina) with custom ZIKV probes. The set included 866 unique probes each of which was 80 nt in length (Supplementary Table 2a). The probes were designed to cover the entire ZIKV genome and to encompass the genetic diversity present on GenBank on 14 January 2016. In total, 26 ZIKV sequences were used during probe design (Supplementary Table 2b). Extracted RNA was fragmented at 94 °C for 0–60 s and each sample was enriched separately using a quarter of the reagents specified in the manufacturer's protocol. Samples were barcoded, pooled and sequenced using the MiSeq Reagent kit v3 (Illumina) on an Illumina MiSeq with a minimum of 2 × 151-bp reads. Dual indexing, with no overlapping indices, was used.

The random hexamer associated with read one and the Illumina adaptors were removed from the sequencing reads using Cutadapt v1.9.dev1 (ref. 50), and low-quality reads or bases were filtered using Prinseq-lite v0.20.3 (ref. 51). Reads were aligned to a reference genome (GenBank: KX197192.1) using Bowtie2 v2.0.6 (ref. 52), duplicates were removed with Picard (http://broadinstitute.github.io/picard), and a new consensus

was generated using a combination of Samtools v0.1.18 (ref. 47) and custom scripts (https://github.com/jtladner/Scripts/blob/master/reference-based_assembly/consensus_fasta.py). Only bases with Phred quality score ≥20 were used in consensus calling, and a minimum of 3 × read-depth coverage, in support of the consensus, was required to make a call; positions lacking this depth of coverage were treated as missing (that is, called as 'N').

Validation and comparison of sequencing methods. The consensus ZIKV sequences from FL01M and FL03M generated by sequencing 35 × 400-bp amplicons on the MiSeq were validated using the following approaches: (1) sequencing the 35 × 400-bp amplicons on the Ion S5 platform (ThermoFisher); (2) sequencing amplicons generated using an Ion AmpliSeq (ThermoFisher) panel custom-targeted towards ZIKV on the Ion S5 platform; and (3) sequencing 5 × 2,150–2,400-bp ZIKV amplicons on the MiSeq. For Ion library preparation, cDNA was synthesized using the SuperScript VILO kit (ThermoFisher). ThermoFisher designed 875 custom ZIKV primers to produce 75 amplicons of ~200 bp in two PCR reactions for use with their Ion AmpliSeq Library Kit 2.0. The reagent FuPa was used to digest the modified primer sequences after amplification. The DNA templates were loaded onto Ion 520 chips using the Ion Chef and sequenced on the Ion S5 with the 200-bp output (ThermoFisher). The 35 × 400-bp amplicons generated for the MiSeq as described above were introduced into the Ion workflow using the Ion AmpliSeq Library Kit 2.0, but without fragmentation. Primers to amplify 2,150–2,400-bp ZIKV fragments (Supplementary Table 2c) were kindly provided by S. O'Connor, D. Dudly, D. O'Connor, and D. Gellerup (AIDS Vaccine Research Laboratory, University of Wisconsin, Madison). Each fragment was amplified individually by PCR using the cDNA generated above, Q5 DNA High-fidelity Polymerase, and the following thermocycle conditions: 55 °C for 30 min, 94 °C for 2 min, 35 cycles of 94 °C for 15 s, 56 °C for 30 s, and 68 °C for 3.5 min, 68 °C for 10 min, and held at 4 °C until use. Each PCR product was purified using Agencourt AMPure XP beads, sheared to 300–400-nt fragments using the Covaris S2 sonicator, indexed and prepared for sequencing as described above, and sequenced using the MiSeq V2 500 cycle kit (paired-end 251-nt reads). Compared to the consensus sequences generated using 35 × 400-bp amplicons on the MiSeq, there were no consensus-level mismatches in the coding sequence using any of the other three approaches (Extended Data Table 2). There were, however, some mismatches in the 5′ and 3′ UTRs (where the genomic RNA is heavily structured), probably as a result of PCR bias and decreased coverage depth.

At least 95% of the ZIKV genome was covered from samples with as low as 4 and 9 GE per μl RNA from the amplicon and enrichment approaches, respectively. These results are similar to our previously determined clinical range of 10–16 ZIKV GE per μl RNA to achieve at least 95% genome coverage using our amplicon-based approach[24]. On average, the amplicon-based sequencing approach covered 97% of the ZIKV genome (≥3 × read-depth) and the targeted enrichment approach covered 82% of the ZIKV genome from clinical samples (Supplementary Table 2d).

Phylogenetic analyses. All published and available complete ZIKV genomes of the Asian genotype from the Pacific and the Americas were retrieved from GenBank public database in December 2016. Public sequences ($n = 65$) were codon-aligned together with ZIKV genomes generated in this study ($n = 39$) using MAFFT[53] and inspected manually. The multiple alignment contained 104 ZIKV sequences collected between 2013 and 2016, from the Pacific (American Samoa, French Polynesia, and Tonga), Brazil, other South and Central Americas (Guatemala, Mexico, Suriname, and Venezuela), the Caribbean (Dominican Republic, Guadeloupe, Haiti, Martinique, and Puerto Rico), and the United States (Supplementary File 1).

To determine the temporal signal of the sequence dataset, a maximum likelihood (ML) phylogeny was first reconstructed with PhyML[54] using the general time-reversible (GTR) nucleotide substitution model and gamma distributed rates amongst sites[55] (Supplementary File 1), which was identified as the best fitting model for ML inference by jModelTest2[56]. Then, a correlation between root-to-tip genetic divergence and date of sampling was conducted in TempEst[57].

Bayesian phylogenetic analyses were performed using BEAST v.1.8.4 (ref. 25) to infer time-structured phylogenies. We used an SDR06 nucleotide substitution model[58] with a non-informative continuous time Markov chain reference prior (CTMC)[59] on the molecular clock rate. Replicate analyses using multiple combinations of molecular clock and coalescent models were explored to select the best fitting model by marginal likelihood comparison using path-sampling and stepping-stone estimation approaches[60–62] (Extended Data Table 1b). The best fit model was a relaxed molecular clock along with a Bayesian Skyline model[63]. All the Bayesian analyses were run for 30 million Markov chain Monte Carlo steps, sampling parameters and trees every 3,000 generations (BEAST XML file and MCC tree available in Supplementary File 1). Support values for all nodes are embedded in the phylogenetic tree files (Supplementary File 1). Tree visualizations were generated with baltic (https://github.com/blab/baltic).

The travel-associated ZIKV genomes add to the Caribbean dataset, but do not directly influence our conclusions about the source of ZIKV introductions into Florida.

Expected number and distribution of local cases from Zika virus importations. We used branching process theory[64,65] to generate the offspring distribution (subsequent local cases) that is expected from a single introduction. The offspring distribution L is modelled with a negative binomial distribution with mean R_0 and over-dispersion parameter k. The total number of cases j that is caused by a single importation (including the index case) after an infinite time[66] has the following form:

$$L = \frac{\Gamma(kj + j - 1)}{\Gamma(kj)\,\Gamma(j+1)} \frac{\left(\frac{R_0}{k}\right)^{j-1}}{\left(1 + \frac{R_0}{k}\right)^{kj+j-1}}$$

The parameter k represents the variation in the number of secondary cases generated by each case of ZIKV[64]. In the case of vector-borne diseases, local heterogeneity is high owing to a variety of factors such as mosquito population abundance, human–mosquito interactions, and control interventions[67–72]. Here, we assumed high heterogeneity ($k = 0.1$) following previous estimates for vector-borne diseases[65]. This distribution L is plotted in Extended Data Fig. 4a. For the following, we took a forward simulation approach, drawing random samples from this distribution. All estimates were based on 100,000 random simulations.

We used this formula to estimate the probability of observing 241 local cases in Miami-Dade County alongside 320 travel-associated cases. We approached this by sampling 320 introduction events from L and calculating the total number of local cases in the resulting outbreak (Extended Data Fig. 4b). We also calculated the likelihood of observing 241 local cases in the total outbreak (Extended Data Fig. 4c), finding that the MLE of R_0 lies between 0.35 and 0.55. As a sensitivity analysis, we additionally modelled introductions with the assumption that only 50% of travellers were infectious at time of arrival into Miami-Dade County, resulting in an MLE of R_0 of 0.45–0.8.

We further used this formula to address the probability of observing 3 distinct genetic clusters (F1, F2 and F3) representing three introduction events in a sample of 27 ZIKV genomes from Miami-Dade County. We approached this by sampling introduction events until we accumulated 241 local cases according to L, arriving at N introduction events with case counts $(j_1, j_2, \ldots j_N)$. We then sampled 27 cases without replacement from $(j_1, j_2, \ldots j_N)$ following a hypergeometric distribution and recorded the number of distinct clusters drawn in the sample. We found that higher values of R_0 resulted in fewer distinct clusters within the sample of 27 genomes (Extended Data Fig. 4d). We additionally calculated the likelihood of sampling three distinct genetic clusters in 27 genomes (Extended Data Fig. 4e), finding an MLE estimate of R_0 of 0.7–0.9. Additionally, as a sensitivity analysis we modelled a preferential sampling process in which larger clusters are more likely to be drawn from than smaller clusters. Here, we used a parameter α that enriches the hypergeometric distribution following $(j_1^\alpha, j_2^\alpha, \ldots j_N^\alpha)$. In this case, we found an MLE estimate of R_0 of 0.5–0.9.

Using the overlap of estimates of R_0 from local case counts (0.35–0.8) and genetic clusters (0.5–0.9), we arrived at a 95% uncertainty range of R_0 of 0.5–0.8. As an additional sensitivity analysis, we incorporated under-reporting in which either 50% of travel-associated cases and 25% of local cases are reported or in which 10% of travel-associated cases and 5% of local cases are reported. We found that differential reporting of travel and local cases resulted in increased mean R_0 estimates when comparing counts of travel-associated to local cases (Extended Data Fig. 4f, g). Additionally, under-reporting increased estimates of R_0 from the sampling analysis (Extended Data Fig. 4h, i). Thus, moderate under-reporting is consistent with R_0 estimates of ~0.8.

We additionally performed birth–death stochastic simulations assuming a serial interval with mean of 20 days[15]. We recorded the number of stochastic simulations still persisting after a particular number of days for different values of R_0 (Fig. 2c).

Zika virus incidence rates. Weekly suspected and confirmed ZIKV case counts from countries and territories within the Americas with local transmission (1 January to 18 September 2016) were obtained from the Pan American Health Organization (PAHO)[30]. In most cases, the weekly case numbers per country were reported only in bar graphs. We contacted PAHO multiple times with the hope of gaining access to the raw data included in the bar graphs, but our requests were unfortunately denied. Therefore we used WebPlotDigitizer v3.10 (http://arohatgi.info/WebPlotDigitizer) to estimate the numbers. We compared the actual ZIKV case numbers reported in Ecuador[73] (the only country with available raw data and reported cases over 10 per week) to our estimates from the PAHO bar graphs and found that the WebPlotDigitizer was ~99% accurate (Extended Data Fig. 5a, b).

Country and territory total population sizes to calculate weekly and monthly ZIKV incidence rates were also obtained from PAHO[74]. Incidence rates calculated from countries and territories in the Americas during January–June 2016 (based on the earliest introduction time estimates until the first known cases) were used as an estimate for infection likelihood to investigate sources of ZIKV introductions.

Airline and cruise ship traffic. To investigate whether the transmission of ZIKV in Florida coincided with travel patterns from ZIKV endemic regions, we obtained the number of passengers arriving at airports in Florida via commercial air travel. We collated data for flights arriving at all commercial airports in Florida from countries and territories in the Americas with local ZIKV transmission between January and June 2016 (based on the earliest introduction time estimates until the first known cases, Supplementary Table 1b). The data were obtained from the International Air Transportation Association, which collects data on an estimated 90% of all passenger trips worldwide. Nelson et al.[28] previously reported flight data from 33 countries with ZIKV transmission entering major US airports from October 2014 to September 2015, which we used to assess the potential for ZIKV introductions outside of Florida.

Schedules for cruise ships visiting Miami, Port Canaveral, Port Everglades, Fort Lauderdale, Key West, Jacksonville (all in Florida), Houston, Galveston (both in Texas), Charleston (South Carolina) and New Orleans (Louisiana) ports in the year 2016 were collated from www.cruisett.com and confirmed by cross-referencing ship logs reported by Port of Miami and reported ship schedules from www.miamidade.gov/portmiami/. Scheduled cruise ship capacities were extracted from www.cruisemapper.com. Every country or territory with ZIKV transmission visited by a cruise ship 10 days (the approximate mean time to ZIKV clearance in human blood (that is, the infectious period))[75] before arrival was counted as contributing the ship's capacity worth of passengers to Miami to the month of arrival (Supplementary Table 1b). While the air traffic was based on the reported number of travellers, we estimated the sea traffic by ship capacity. Lee and Ramdeen[76] reported that the average occupancy of cruise ships travelling to the Caribbean Islands exceeded 100% in 2011, and according to the Florida-Caribbean Cruise Association[77], it remained >100% in 2015. Occupancy data for 2016 was not available at the time of publication, but we assumed that it was also near 100%.

Expected number of travellers infected with Zika virus. We estimated the expected number of travellers entering Miami who were infected with ZIKV (λ) by using the total travel capacity (C) and the likelihood of ZIKV infection (infections (I) per person (N)) from each country/territory (i):

$$\lambda = \sum_i C_i \frac{I_i}{N_i}$$

We summed the number of expected infected travellers from each country or territory with ZIKV transmission by region and travel method (flights or cruises). The number of ZIKV cases reported by each country are likely to be underestimates, in part because the majority of ZIKV infections are asymptomatic[2,31]. We normalized some of the potential reporting variances between countries by reporting the data as the relative proportion of infected travellers (Fig. 3c, Extended Data Fig. 7a) and as the absolute number of infected travellers (Fig. 3d, Extended Data Fig. 7b, Supplementary Table 1b) from each region. We also accounted for potential reporting biases with incidence rates by using ZIKV attack rates (that is, proportion infected before epidemic burnout) to estimate peak transmission intensity. Attack rates were calculated using a susceptible–infected–recovered (SIR) transmission model derived from seroprevalence studies and environmental factors as described[78]. Using attack rates as an estimate of infection likelihood, we predict that ~60% of the infected travellers entering Miami came from the Caribbean (Extended Data 7b), which is in agreement with our methods using incidence rates of ~60–70% (Fig. 3c). A list of countries and territories used in these analyses can be found in Supplementary Table 1b.
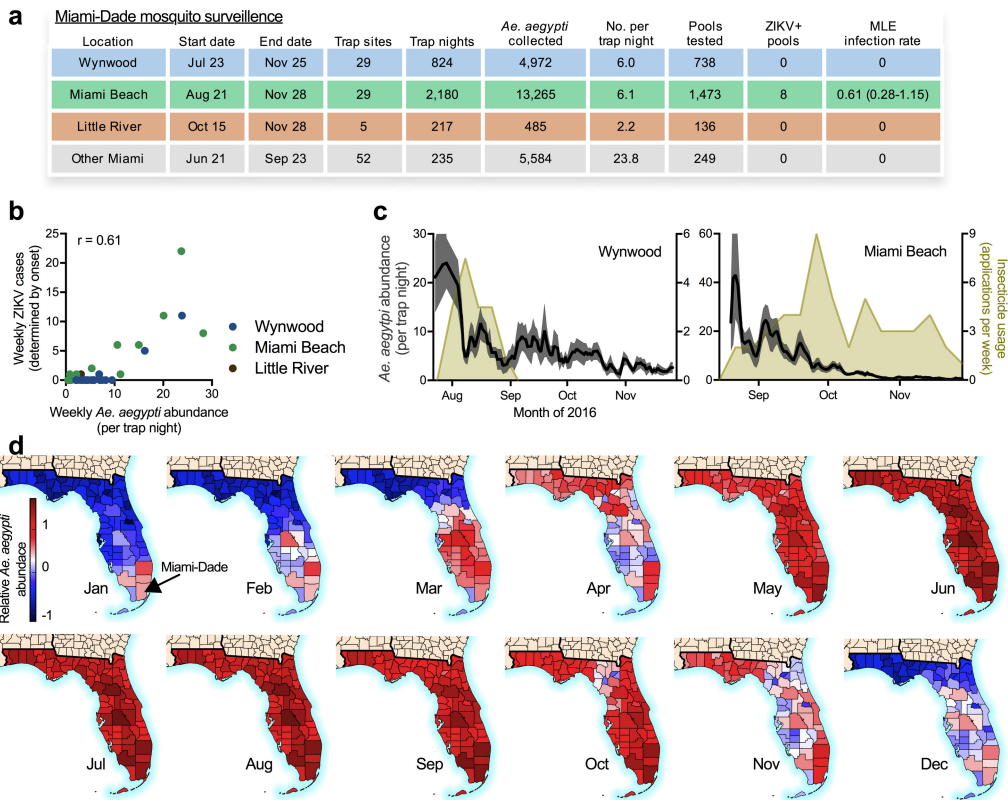
Maps. The maps presented in our figures were generated using Matplotlib[79] and ESRI basemaps (www.esri.com/data/basemaps). The software and basemaps are open source and freely available to anyone.

Data availability. All ZIKV sequencing data are available under NCBI BioProjects PRJNA342539 and PRJNA356429. Individual sample GenBank access numbers are listed in Supplementary Table 1a. All other data are available in the Extended Data or Supplementary Information, or upon request.

34. Zika virus. *Florida Department of Health* http://www.floridahealth.gov/diseases-and-conditions/zika-virus/index.html?utm_source=flhealthIndex (2017).
35. Daily Zika update. *Florida Department of Health* http://www.floridahealth.gov/newsroom/all-articles.html (2017).
36. Zika virus case counts in the US. *The Centers for Disease Control and Prevention* http://www.cdc.gov/zika/geo/united-states.html (2016).
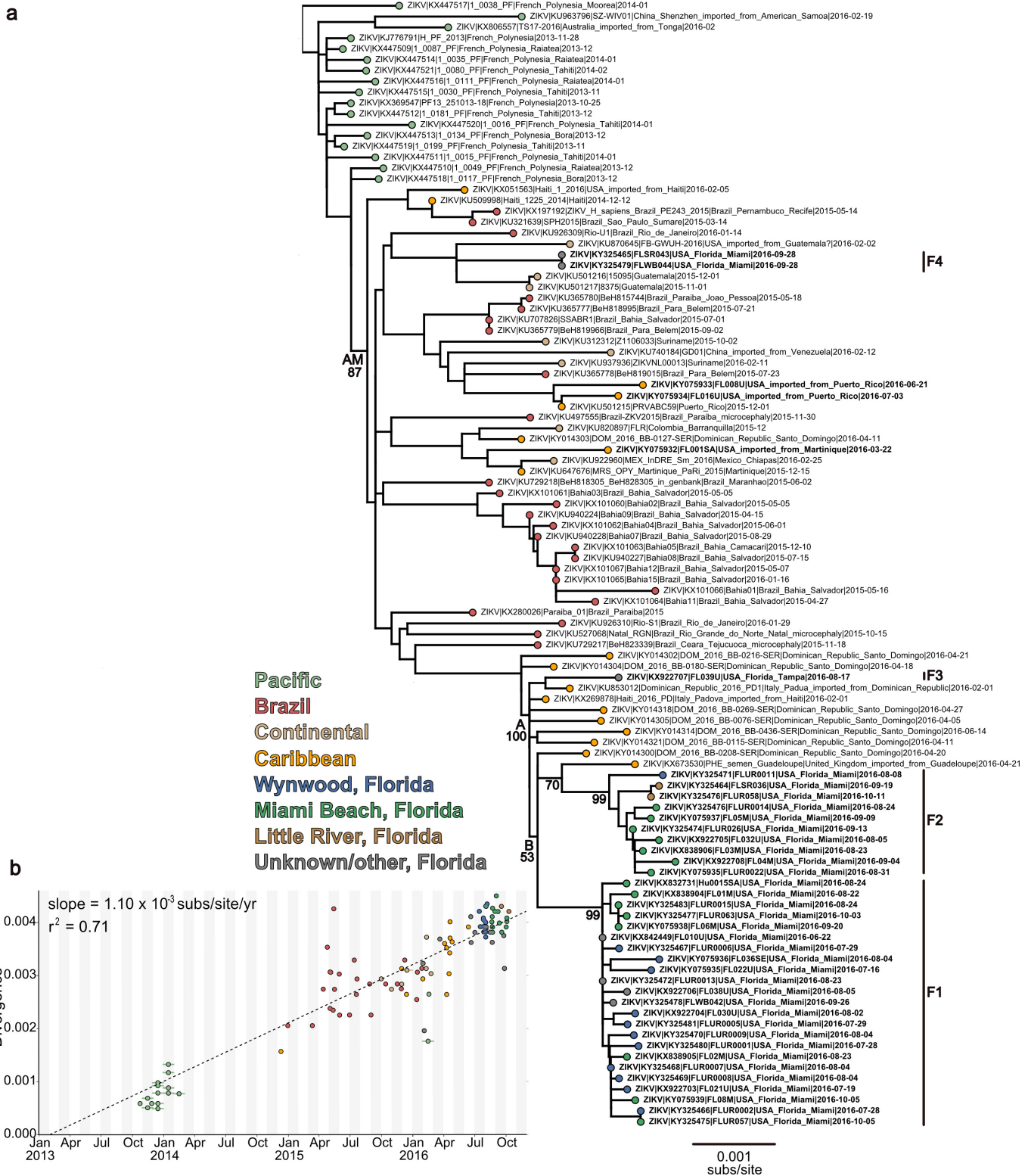
37. Rabe, I. B. et al. Interim guidance for interpretation of Zika virus antibody test results. *MMWR Morb. Mortal. Wkly. Rep.* **65,** 543–546 (2016).
38. Waggoner, J. J. & Pinsky, B. A. Zika virus: diagnostics for an emerging pandemic threat. *J. Clin. Microbiol.* **54,** 860–867 (2016).
39. Lanciotti, R. S. et al. Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* **14,** 1232–1239 (2008).
40. Interim Guidance for Zika Virus Testing of Urine—United States, 2016. *MMWR Morb. Mortal. Wkly. Rep.* **65,** 474 (2016).
41. Biggerstaff, B. *PooledInfRate, Version 4.0* (Centers for Disease Control, 2009).
42. Bogoch, I. I. et al. Potential for Zika virus introduction and transmission in resource-limited countries in Africa and the Asia-Pacific region: a modelling study. *Lancet Infect. Dis.* **16,** 1237–1245 (2016).
43. Hwang, W.-H. & He, F. Estimating abundance from presence/absence maps. *Methods Ecol. Evol.* **2,** 550–559 (2011).
44. Corman, V. M. et al. Clinical comparison, standardization and optimization of Zika virus molecular detection. *Bull. World Health Organ.* http://dx.doi.org/10.2471/BLT.16.175950 (2016).
45. Worobey, M. et al. 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature* **539,** 98–101 (2016).
46. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).
47. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
48. Köster, J. & Rahmann, S. in *OASIcs-OpenAccess Series in Informatics* **26,** (Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012).
49. Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28,** 1647–1649 (2012).
50. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).
51. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27,** 863–864 (2011).
52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
53. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772–780 (2013).
54. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704 (2003).
55. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39,** 306–314 (1994).
56. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9,** 772 (2012).
57. 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.* http://dx.doi.org/10.1093/ve/vew007 (2016).
58. Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23,** 7–9 (2006).
59. Ferreira, M. A. R. & Suchard, M. A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **36,** 355–368 (2008).
60. Baele, G. et al. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29,** 2157–2167 (2012).
61. Xie, W., Lewis, P. O., Fan, Y., Kuo, L. & Chen, M.-H. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* **60,** 150–160 (2011).
62. Gelman, A. & Meng, X.-L. Simulating normalizing constants: from importance sampling to bridge sampling to path sampling. *Stat. Sci.* **13,** 163–185 (1998).
63. Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22,** 1185–1192 (2005).
64. Churcher, T. S. et al. Public health. Measuring the path toward malaria elimination. *Science* **344,** 1230–1232 (2014).
65. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438,** 355–359 (2005).
66. Nishiura, H., Yan, P., Sleeman, C. K. & Mode, C. J. Estimating the transmission potential of supercritical processes based on the final size distribution of minor outbreaks. *J. Theor. Biol.* **294,** 48–55 (2012).
67. Perkins, T. A., Scott, T. W., Le Menach, A. & Smith, D. L. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. *PLOS Comput. Biol.* **9,** e1003327 (2013).
68. Kraemer, M. U. G. et al. Big city, small world: density, contact rates, and transmission of dengue across Pakistan. *J. R. Soc. Interface* **12,** 20150468 (2015).
69. Kraemer, M. U. G. et al. Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. *Lancet Infect. Dis.* **17,** 330–338 (2016).
70. Struchiner, C. J., Rocklöv, J., Wilder-Smith, A. & Massad, E. Increasing dengue incidence in Singapore over the past 40 years: population growth, climate and mobility. *PLoS One* **10,** e0136286 (2015).

71. Fauver, J. R. *et al.* Temporal and spatial variability of entomological risk indices for West Nile virus infection in Northern Colorado: 2006–2013. *J. Med. Entomol.* **53,** 425–434 (2016).
72. Rocklöv, J. *et al.* Assessing seasonal risks for the introduction and mosquito-borne spread of Zika virus in Europe. *EBioMedicine* **9,** 250–256 (2016).
73. National Direction of Epidemiological Surveillance for Vector Transmitted Diseases (Spanish) *Ministry of Public Health, Ecuador*. http://www.salud.gob.ec/wp-content/uploads/2015/12/GACETA-ZIKA_SE5corregido.pdf (2017).
74. Zika Cumulative Cases. *Pan American Health Organization* http://www.paho.org/hq/index.php?option=com_content&view=article&id=12390&Itemid=42090&lang=en (2016).
75. Lessler, J. T. *et al.* Times to key events in the course of Zika infection and their implications: a systematic review and pooled analysis. *Bull. World Health Organ.* **94,** 841–849 (2016).
76. Lee, S. & Ramdeen, C. Cruise ship itineraries and occupancy rates. *Tourism Manage.* **34,** 236–237 (2013).
77. FCCA Research & Statistics. *Florida-Caribbean Cruise Association* http://www.f-cca.com/research.html (2017).
78. Alex Perkins, T., Siraj, A. S., Ruktanonchai, C. W., Kraemer, M. U. G. & Tatem, A. J. Model-based projections of Zika virus infections in childbearing women in the Americas. *Nat. Microbiol.* **1,** 16126 (2016).
79. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9,** 90–95 (2007).

a Miami-Dade mosquito surveillence

Location	Start date	End date	Trap sites	Trap nights	*Ae. aegypti* collected	No. per trap night	Pools tested	ZIKV+ pools	MLE infection rate
Wynwood	Jul 23	Nov 25	29	824	4,972	6.0	738	0	0
Miami Beach	Aug 21	Nov 28	29	2,180	13,265	6.1	1,473	8	0.61 (0.28-1.15)
Little River	Oct 15	Nov 28	5	217	485	2.2	136	0	0
Other Miami	Jun 21	Sep 23	52	235	5,584	23.8	249	0	0

Extended Data Figure 1 | Miami-Dade mosquito surveillance and relative *A. aegypti* abundance. a, Mosquito surveillance data reported from 21 June to 28 November 2016 was used to evaluate the risk of ZIKV infection from mosquito-borne transmission in Miami. A total of 24,306 *A. aegypti* and 45 *A. albopictus* were collected. Trap nights are the total number of times each trap site was used and the trap locations are shown in Fig. 1d (some 'Other Miami' trap sites are located outside the mapped region). Up to 50 mosquitoes of the same species and trap night were pooled together for ZIKV RNA testing. The infection rates were calculated using an MLE. None of the *A. albopictus* pools contained ZIKV RNA. **b**, The number of weekly ZIKV cases (based on symptom onset) was

correlated with mean *A. aegypti* abundance per trap night determined from the same week and zone (Spearman $r = 0.61$). This suggests that when the virus is present, mosquito abundance numbers alone could be used to target control efforts. **c**, Insecticide usage, including truck and aerial adulticides and larvacides, by Miami-Dade Mosquito Control in Wynwood (left) and Miami Beach (right) was overlaid with *A. aegypti* abundance per trap night to demonstrate that intense usage of insecticides may have helped to reduce local mosquito populations. **d**, Relative *A. aegypti* abundance for each Florida county and month was estimated using a multivariate regression model, demonstrating spatial and temporal heterogeneity for the risk of ZIKV infection.
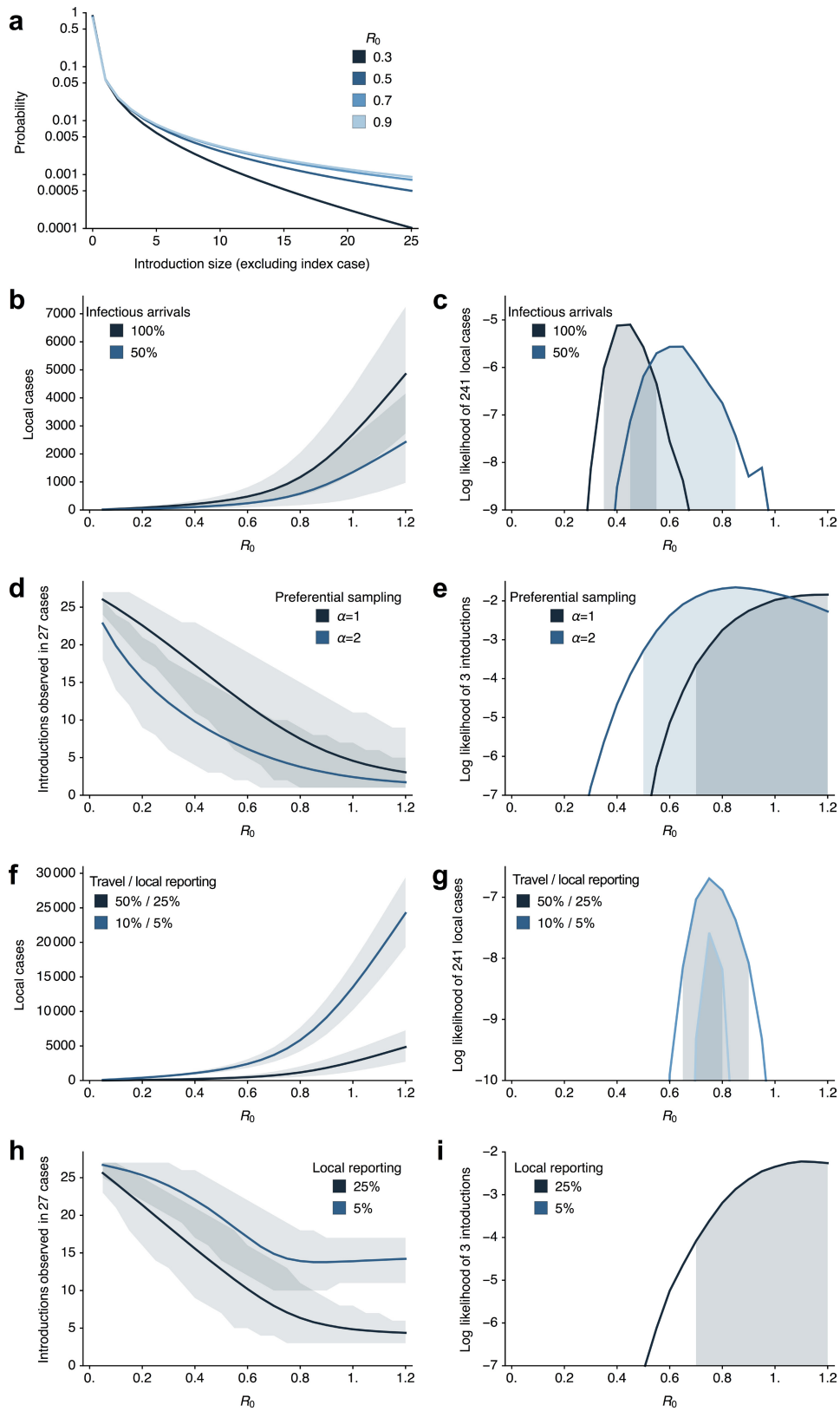
Extended Data Figure 2 | Maximum likelihood tree and root-to-tip regression of Zika virus genomes from Pacific islands and the epidemic in the Americas. a, Maximum likelihood tree of publicly available ZIKV sequences and sequences generated in this study ($n = 104$). Tips are coloured by location and labels in bold indicate sequences generated in this study. Florida clusters F1–F4 are indicated by vertical lines to the right of the tree. Bootstrap support values are shown at key nodes. All other support values can be found in Supplementary File 1. **b**, Linear regression of sample tip dates against divergence from root based on sequences

with known collection dates estimates an evolutionary rate for the ZIKV phylogeny of 1.10×10^{-3} nucleotide substitutions per site per year (subs/site/yr). This is consistent with BEAST analyses using a relaxed molecular clock and a Bayesian Skyline tree prior, the best performing combination of clock and demographic model according to MLEs (Extended Data Table 1c), which estimated an evolutionary rate of 1.21×10^{-3} (95% highest posterior density: $1.01–1.43 \times 10^{-3}$) substitutions per site per year (Extended Data Table 1a). These values are in agreement with previous estimates based on ZIKV genomes from Brazil[6].

Extended Data Figure 3 | Molecular clock dating of Zika virus clades.
Maximum clade credibility (MCC) tree of ZIKV genomes collected from
Pacific islands and the epidemic in Americas ($n = 104$). Circles at the tips
are coloured according to origin location. Clade posterior probabilities
are indicated by white circles filled with black relative to the support.
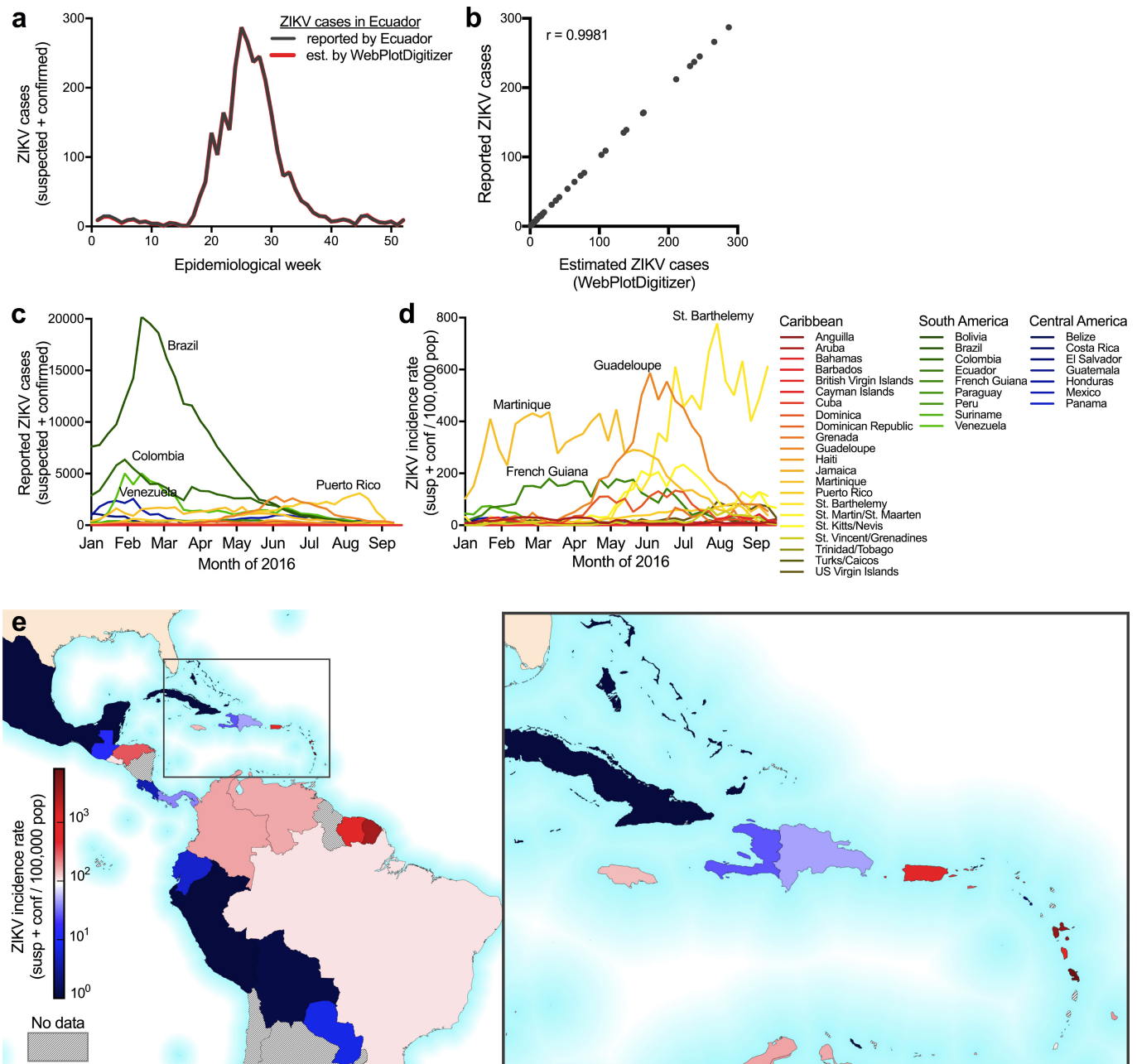A posterior probability of 1 fills the entire circle black. The grey violin plot
indicates the 95% HPD interval for the tMRCA of the American epidemic.
We estimated that the tMRCA for the ongoing epidemic in the Americas
occurred during October 2013 (node AM, Extended Data Table 1, 95%
HPD: August 2013–January 2014), which is consistent with previous
analysis based on ZIKV genomes from Brazil[6].
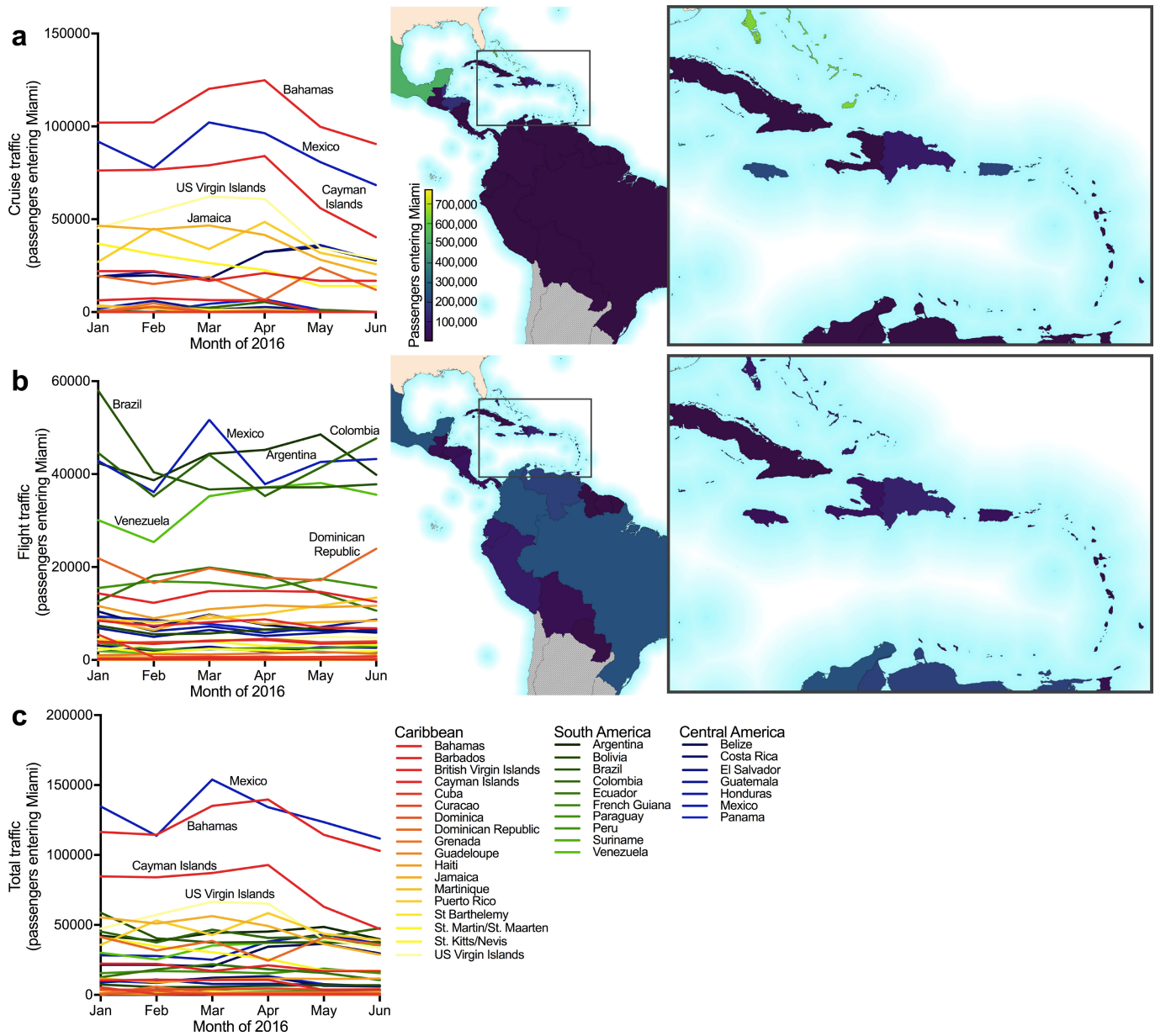
Extended Data Figure 4 | See next page for caption.

Extended Data Figure 4 | Estimation of basic reproductive number and number of introductions in Miami-Dade County. a, Probability distribution of estimated total number of cases caused by a single introduction (excluding the index case) for different values of R_0. **b**, Mean and 95% CI for total number of local cases caused by 320 introduction events (that is, travel-associated cases diagnosed in Miami-Dade County) for different values of R_0 and for different assumptions of proportion of infectious travellers. **c**, log likelihood of observing 241 local cases in Miami-Dade County with 320 introduction events for different values of R_0 along with 95% MLE bounds on R_0. **d**, Mean and 95% uncertainty interval for total number of distinct phylogenetic clusters observed in 27 sequenced ZIKV genomes from human cases diagnosed in Miami-Dade County for different values of R_0 and for different assumptions of sampling bias, from $\alpha = 1$ (no sampling bias) to $\alpha = 2$ (skewed towards preferentially sampling larger clusters). **e**, log likelihood of observing three clusters (that is, ZIKV lineages F1, F2, and F4, Fig. 2a) in 27 sequenced

cases for different values of R_0 along with 95% MLE bounds on R_0. **f**, Mean and 95% CI for total number of local cases caused by 320 observed travel-associated cases with travel-associated versus local reporting rates of 50%/25% and 10%/5%. This assumes that 50% of travellers are infectious. **g**, log likelihood of observing 241 local cases with 320 introduction events for different values of R_0 along with 95% MLE bounds on R_0 with travel-associated versus local reporting rates of 50%/25% and 10%/5%. **h**, Mean and 95% uncertainty interval for total number of distinct phylogenetic clusters observed in 27 sequenced ZIKV genomes for different values of R_0 and for assumptions of local reporting rate of 5% and 25%. This assumes preferential sampling ($\alpha = 2$). **i**, log likelihood of observing three clusters in 27 sequenced cases for different values of R_0 along with 95% MLE bounds on R_0 with local reporting rates of 5% and 25%. At 5% local reporting rate, none of the 100,000 replicates for all R_0 values showed three clusters.
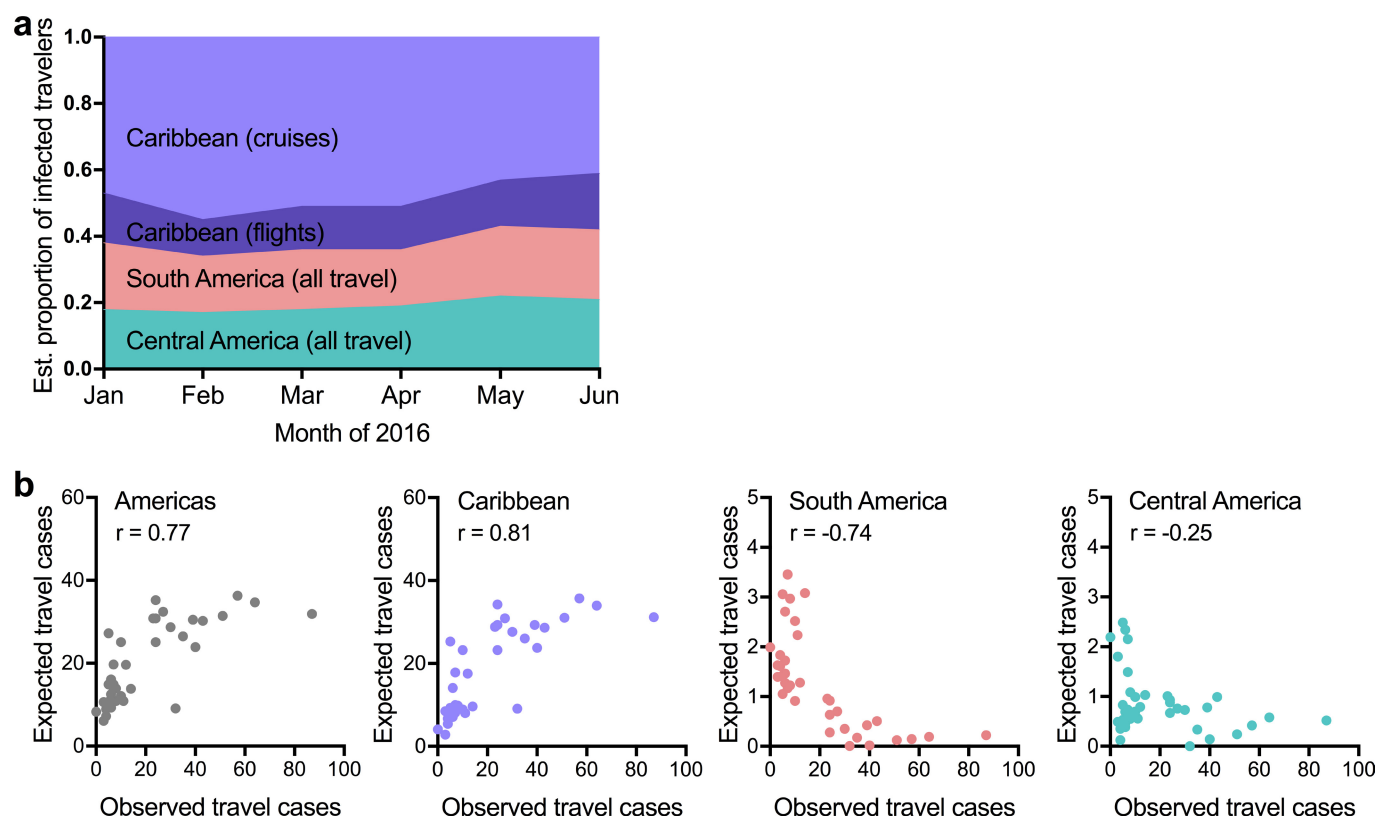
Extended Data Figure 5 | Weekly reported Zika virus case numbers and incidence rates in the Americas. a, Most ZIKV case numbers reported by PAHO[30] were available only as bar graphs (raw data were not made available to us at the time of request). Therefore we used WebPlotDigitizer to estimate the weekly case numbers from the PAHO bar graphs. ZIKV cases reported from Ecuador was the only dataset to include a link to the actual case numbers that also had more than 10 cases per week[73]. To validate the WebPlotDigitizer-derived values, we compared the weekly reported case numbers from Ecuador to our estimates. **b,** The reported and estimated case numbers were strongly correlated (Spearman $r = 0.9981$). WebPlotDigitizer was used to estimate the ZIKV case numbers for all subsequent analysis. **c, d,** ZIKV cases (suspected and confirmed; **c**) and incidence rates (normalized per 100,000 population; **d**) are shown for each country or territory with available data per epidemiological week from 1 January to 18 September 2016. **e,** Each country or territory with available data is coloured by its reported ZIKV incidence rate from January to June 2016 (the time frame for analysis of ZIKV introductions into Florida).

Extended Data Figure 6 | Cruise and flight traffic entering Miami from regions with Zika virus transmission. a, b, The estimated number of passengers entering Miami, by cruises (**a**) or flights (**b**), from each country or territory in the Americas with ZIKV transmission per month (left). The centre map and inset show the cumulative numbers of travellers entering Miami during January to June 2016 (the time frame for analysis of ZIKV introductions into Florida) from each country or territory per method of travel. **c,** The total traffic (that is, cruises and flights) entering Miami per month.

Extended Data Figure 7 | Expected number of ZIKV-infected travellers from the Caribbean correlated with the total observed number of travel-associated infections. a, To account for potential biases in ZIKV reporting accuracy, we also estimated the proportion of infected travellers using projected ZIKV attack rates[78] (that is, predicted proportion of population infected before epidemic burnout). About 60% of the infected travellers are expected to have arrived from the Caribbean, similar to our results using incidence rates (Fig. 3c). **b**, The expected number of travel-associated ZIKV cases was estimated by the number of travellers coming into Miami from each country or territory (travel capacity) and the in-country or in-territory infection likelihood (incidence rate per person) per week. The expected travel cases were summed from all of the Americas (left), Caribbean (left centre), South America (right centre), and Central America (right) and plotted with the observed travel-associated ZIKV cases. Numbers in each plot indicate Spearman correlation coefficients. Negative Spearman *r* coefficients indicate a negative correlation between the number of expected and observed travel cases.

Extended Data Figure 8 | Greater early season potential for Zika virus introductions into Miami. The monthly cruise ship and airline[28] capacity from countries or territories with ZIKV transmission for the major US travel hubs (shown as circle diameter) with monthly potential *A. aegypti* abundance (circle colour), as previously estimated[22]. The abundance ranges were chosen with respect to the May–October Miami mean: None to low (<2%); Low to moderate (2–25%); Moderate to high (25–75%); and High (>75%). Mosquito-borne transmission is unlikely in the 'None to low' range. Cruise capacities from Houston and Galveston, Texas were combined.

Extended Data Table 1 | Evolutionary analyses and model selection

a

Model combination	Clade AM tMRCA			Clade A tMRCA			Clade B tMRCA		
	Mean	Lower 95% HPD	Higher 95% HPD	Mean	Lower 95% HPD	Higher 95% HPD	Mean	Lower 95% HPD	Higher 95% HPD
Strict, Constant	2013.87	2013.65	2014.09	2015.59	2015.43	2015.76	2015.70	2015.52	2015.86
Strict, Exponential	2013.90	2013.69	2014.09	2015.55	2015.36	2015.72	2015.66	2015.47	2015.83
Strict, Bayesian Skyline	2013.92	2013.70	2014.12	2015.56	2015.38	2015.74	2015.69	2015.50	2015.86
UCLN, Constant	2013.95	2013.66	2014.21	2015.70	2015.49	2015.88	2015.82	2015.63	2016.03
UCLN, Exponential	2013.95	2013.70	2014.19	2015.62	2015.39	2015.81	2015.73	2015.51	2015.93
UCLN, Bayesian Skyline	2013.97	2013.72	2014.23	2015.62	2015.40	2015.81	2015.76	2015.54	2015.94

Model combination	Clade F1 tMRCA			Clade F2 tMRCA			Evolutionary Rate		
	Mean	Lower 95% HPD	Higher 95% HPD	Mean	Lower 95% HPD	Higher 95% HPD	Mean	Lower 95% HPD	Higher 95% HPD
Strict, Constant	2016.16	2016.02	2016.28	2016.16	2016.00	2016.31	1.09E-03	9.27E-04	1.23E-03
Strict, Exponential	2016.12	2015.97	2016.25	2016.13	2015.94	2016.28	1.08E-03	9.35E-04	1.23E-03
Strict, Bayesian Skyline	2016.25	2016.13	2016.36	2016.24	2016.06	2016.37	1.13E-03	9.80E-04	1.28E-03
UCLN, Constant	2016.19	2016.05	2016.31	2016.21	2016.05	2016.36	1.22E-03	9.95E-04	1.47E-03
UCLN, Exponential	2016.14	2015.98	2016.27	2016.16	2015.98	2016.32	1.17E-03	9.68E-04	1.39E-03
UCLN, Bayesian Skyline	2016.27	2016.16	2016.38	2016.27	2016.08	2016.39	1.21E-03	1.01E-03	1.43E-03

b

Model combination	Path Sampling	Ranking	Stepping Stone	Ranking
Strict, Constant	-20379.871	5	-20384.201	5
Strict, Exponential	-20375.747	4	-20379.272	4
Strict, Bayesian Skyline	-20358.676	2	-20363.748	2
UCLN, Constant	-20381.228	6	-20385.807	6
UCLN, Exponential	-20370.555	3	-20374.418	3
UCLN, Bayesian Skyline	-20354.940	1	-20361.218	1

a, Time of the most recent common ancestor and evolutionary rate. b, Model selection to infer time-structured phylogenies. HPD, highest posterior density. Dates listed as proportion of days elapsed with a year. Clades refer to Fig. 2a.

Extended Data Table 2 | Validation of sequencing results

Sample	Amplicon method	NGS platform	Mismatches/nucleotides covered[a]		
			3' UTR	CDS	5' UTR
FL01M	35 × 400 bp	Ion S5	1/80	0/10272	7/252
	75 × ~200 bp[b]	Ion S5	2/75	0/10272	4/205
	5 × ~2,200 bp	MiSeq	0/80	0/10272	0/32
FL03M	35 × 400 bp	Ion S5	3/87	0/10272	20/252
	75 × ~200 bp[b]	Ion S5	4/78	0/10272	5/198
	5 × ~2,200 bp	MiSeq	0/82	0/10272	0/32

NGS, next-generation sequencing; UTR, untranslated region; CDS, coding sequence.
[a]Compared to the consensus genomes generated by sequencing 35 × 400-bp amplicons on the MiSeq.
[b]Amplicons produced using Ion AmpliSeq and 875 custom ZIKV primers.