Emergence of Novel Norovirus GII.4 Variant

Preeti Chhabra,[1] Damien C. Tully,[1] Janet Mans, Sandra Niendorf, Leslie Barclay,
Jennifer L. Cannon, Anna M. Montmayeur, Chao-Yang Pan, Nicola Page, Rachel Williams,
Helena Tutill, Sunando Roy, Cristina Celma, Stuart Beard, Michael L. Mallory,
Gédéon Prince Manouana, Thirumalaisamy P. Velavan, Ayola Akim Adegnika,
Peter G. Kremsner, Lisa C. Lindesmith, Stéphane Hué, Ralph S. Baric, Judith Breuer, Jan Vinjé

We detected a novel GII.4 variant with an amino acid insertion at the start of epitope A in viral protein 1 of noroviruses from the United States, Gabon, South Africa, and the United Kingdom collected during 2017–2022. Early identification of GII.4 variants is crucial for assessing pandemic potential and informing vaccine development.

Norovirus is the most common cause of acute gastroenteritis (AGE) worldwide. Norovirus has an ≈7.7 kb positive-sense single-stranded RNA genome organized into 3 open reading frames (ORFs). ORF1 encodes a polyprotein that is post-translationally cleaved into 6 nonstructural (NS) proteins, including NS7, the viral RNA-dependent RNA polymerase (RdRp). ORF2 encodes the major viral protein (VP), VP1, and ORF3 encodes the minor VP2 capsid protein.

Noroviruses are genetically diverse and classified into ≥10 different genogroups. Genogroup II genotype 4 (GII.4) viruses cause most illnesses worldwide (1,2). GII.4 variants include US95–96, Farmington Hills_2002, Asia_2003, Hunter_2004, Yerseke_2006, Den Haag_2006, Osaka_2007, Apeldoorn_2007, New Orleans_2009, Sydney_2012, and HongKong_2019 (3). GII.4 variant emergence has been associated with changes in epitopes A–I on the surface exposed P2 subdomain of VP1 affecting interactions with histo-blood group antigens (HBGA) on host cells (4). Since 2012, GII.4 Sydney has been the most prevalent norovirus genotype globally (2). We sequenced complete genomes or VP1 of GII.4 viruses from recent outbreaks and sporadic cases that could not be genotyped to investigate genomic similarities with existing variants.

The Study

Several surveillance networks track trends in norovirus strain diversity, including CaliciNet in the United States (5) and NOROPATROL in the United Kingdom. In July 2017, four identical sequences from a norovirus outbreak in a childcare facility in San Francisco, California, USA, were uploaded to CaliciNet (https://www.cdc.gov/norovirus/reporting/calicinet). We genotyped those 4 sequences as GII.4 untypeable; the sequences had >2% nucleotide sequence difference in the 5′ end of ORF2 from existing GII.4 viruses. We identified genetically similar strains in stool specimens from 3 UK outbreaks: 1 strain from a hospitalized 32-year-old patient with AGE in Newcastle in 2019, 1 from

Table. List of oligonucleotide primers used to examine emergence of novel norovirus GII.4 strains on 3 continents*

Primer names	Sequence, 5′ → 3′	Orientation	Nested reverse transcription PCR
NV6e	ACC AYT WTG ATG CAG ACT A	Forward	Round 1
NV6f	ACC AYT ATG ATG CTG ATT A	Forward	Round 1
NV6g	ATC AYT ATG ATG CWG AYT A	Forward	Round 1
NV357a	CGC CAG TCC AGG AGT CCA AAA TY	Reverse	Round 1
NV378	GCT TAC GAA TGT GAG CGA A	Reverse	Round 2
*Primers were used to amplify complete ORF2 and ORF3 genes of GII.4 San Francisco strains from Gabon. ORF, open reading frame			

a 1-year-old child in London in 2021, and 2 from Brighton in 2021, from a 1-year-old and a 3-year-old from the same household in an oyster-related outbreak. We also identified similar strains in sporadic samples from children with AGE in Gabon during 2018–2019 (*6*) and in children and adolescents with AGE from Western Cape and Gauteng in South Africa during 2021–2022. We sequenced 5 near-complete genomes and 10 complete VP1 sequences and compared those with existing GII.4
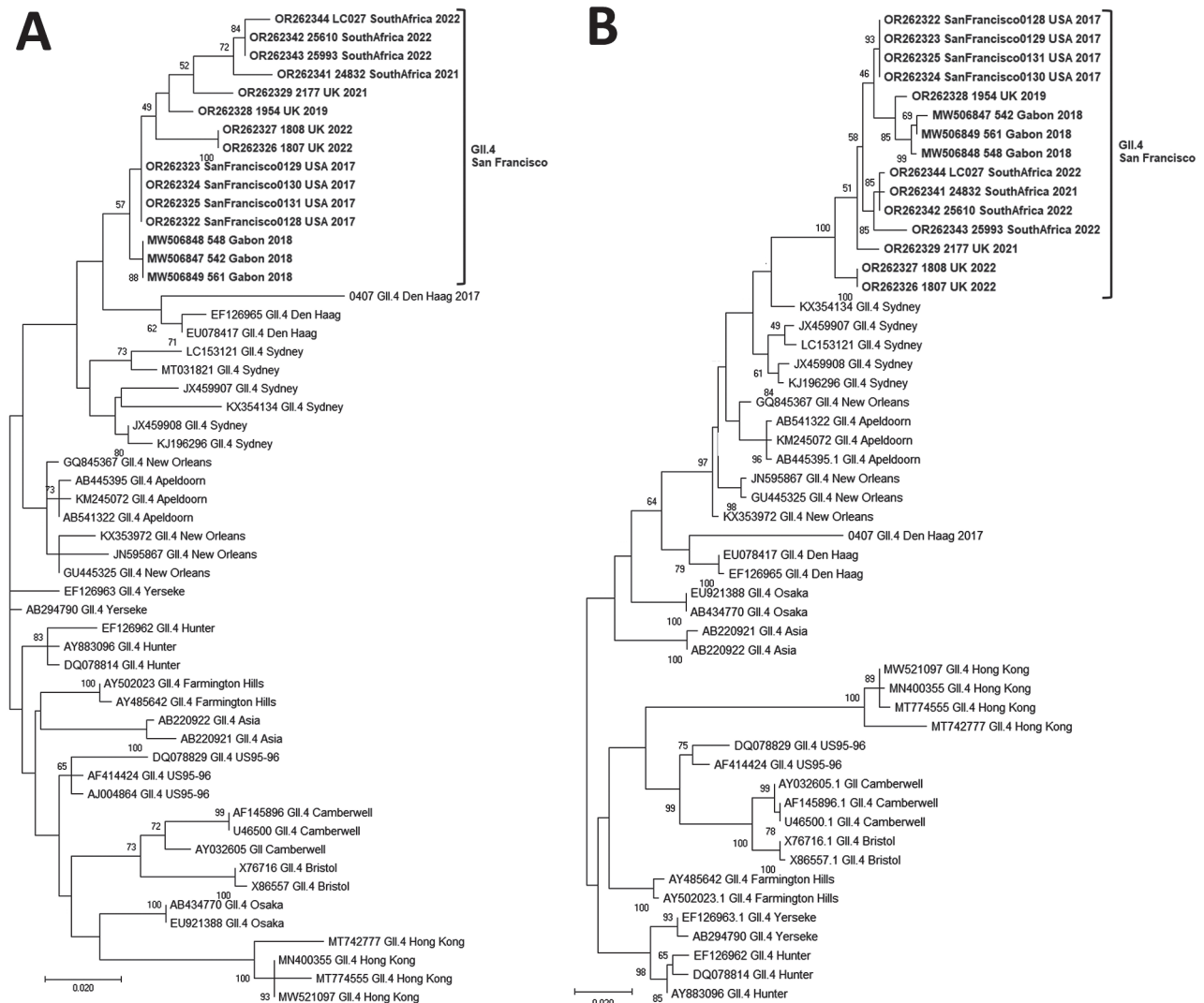

Figure 1. Phylogenetic trees of the emerging novel norovirus GII.4 strains on 3 continents. A) Common genotyping region C; 250 nt from the 5′ end of ORF2; B) complete VP1 aa sequences. Phylogenetic trees show novel GII.4 San Francisco strains and GII.4 variants, including recently identified clusters like GII.4 Hong Kong. Evolutionary analyses were conducted in MEGA X (https://www.megasoftware.net) using the maximum-likelihood method based on the Tamura-Nei model for the C region and Jones-Taylor-Thornton matrix-based model for VP1. We used a discrete gamma distribution to model evolutionary rate differences among sites; 5 categories γ parameter = 0.2174. Bootstrap (100) values are indicated at the nodes. Trees were drawn to scale. Scale bar represents nucleotide substitutions per site. ORF, open reading frame; VP1, viral protein 1.

Figure 2. Structural changes of emergent novel norovirus GII.4 strains from 3 continents. A) Sydney GII.4 strain (GenBank accession no. JX459908); B) GII.4 San Francisco strain. The 3-dimensional structure models were predicted by using ChimeraX version 1.4 (*11*) and the alphafold prediction tool (*12*). Models show structural changes near and within the epitope A antigenic region on GII.4 San Francisco P-domain (panel B) are overlayed on a GII.4 Sydney 2012 backbone (Protein Data Bank, https://www.rcsb.org/ structure/4OP7). Negatively

(red) and positively (green) charged amino acids of GII.4 Sydney (panel A) were replaced with neutral amino acids (blue) in the GII.4 San Francisco strain and a hydrophobic (yellow) amino acid, alanine, was inserted between T293 and T294.
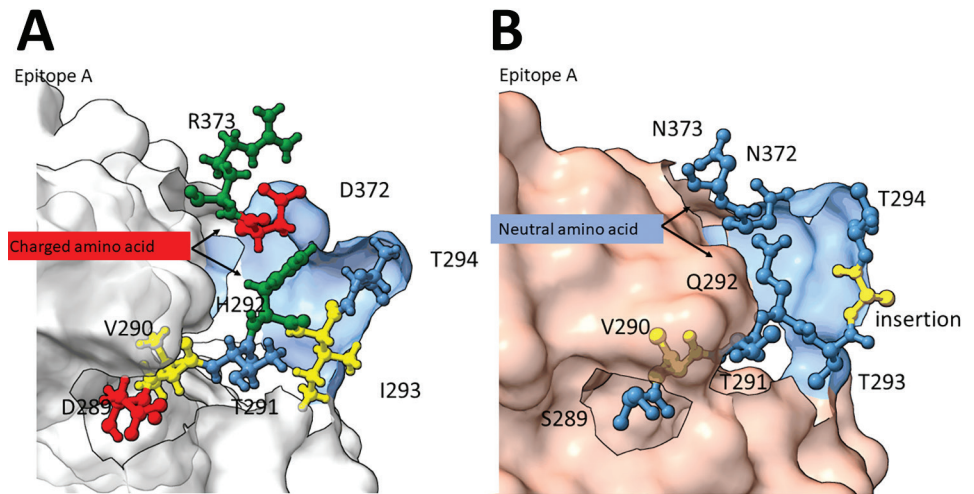
sequences from the Human Calicivirus Typing Tool (https://calicivirustypingtool.cdc.gov/gebali.cgi).

We extracted viral RNA and obtained complete genome or VP1 sequences for strains from the United States, United Kingdom, and South Africa according to published methods (*5–9*). We amplified the complete VP1 from Gabon strains by seminested reverse transcription PCR (RT-PCR) using Oligo dT and Lunascript Master Mix Kit (New England Biolabs, https://www.neb.com) for cDNA synthesis at 55°C for 30 min. We amplified complete VP1 and VP2 by using seminested RT-PCR and oligonucleotide primers designed for this study (Table). We performed RT-PCRs by using OneTaq 2X Master Mix (New England Biolabs) for 30 cycles at 94°C for 10 s, 45°C for 30 s, and 72°C for 3 min, then a final extension of 72°C for 2 min. We sequenced all amplicons.

We aligned complete VP1 amino acid sequences with GII.4 reference strains representing all known emerging and epidemic GII.4 viruses by using ClustalW in MEGA X (*10*). We computed maximum-likelihood phylogenetic trees by using the Jones-Taylor-Thornton model for amino acid sequences and Tamura-Nei model for nucleotide sequences and performed gamma distribution of evolutionary rates among sites using 100 bootstrap replications. We deposited nucleotide sequences of GII.4 San Francisco strains in GenBank (accession nos. OR262322–29, OR262341–44, and MW506847–49). We predicted 3-dimensional structures of GII.4 San Francisco viruses by using ChimeraX version 1.4 (*11*) and the alphafold prediction tool (*12*) and used the P-domain of GII.4 Sydney (Protein Data Bank no. PDB 4OP7;

https://www.rcsb.org/structure/4OP7) as the backbone. To evaluate the effects of amino acid changes in P2, we synthesized virus-like particles from the codon-optimized ORF2 sequence of SF128 (GenBank accession no. OR262322) and compared ligand binding with other GII.4 virus-like particles (*4*).

We found that GII.4 San Francisco sequences from the 5'-end of ORF2 were closest to GII.4 Sydney and GII.4 Den Haag reference strains with maximum identities ranging from 91%–95% (Figure 1, panel A). Complete VP1 amino acid sequences of GII.4 San Francisco strains formed a distinct cluster with 5%–10% amino acid difference from GII.4 New Orleans and GII.4 Sydney (Figure 1, panel B). We typed RdRp sequences of all strains as GII.P31.

Of note, VP1 sequences of all GII.4 San Francisco strains had an alanine insertion at position 293/294 at the start of epitope A, coinciding with a unique SVTQTAT/A motif at positions 289–295 adjacent to epitope A (Appendix Figure 1, https://wwwnc.cdc. gov/EID/article/30/1/23-1003-App1.pdf). Compared with GII.4 Sydney_2012 and GII.4 New Orleans viruses, we observed mutations at amino acid residues 256 and 438 in the P1 region and 294, 310, 340, 341, 356, 372, 373, 377, 393, and 395 in the hypervariable region, P2 (Appendix Figure 1).

Homology modeling of the GII.4 San Francisco P-domain using GII.4 Sydney 2012 as a backbone (PDB 4OP7; GenBank accession no. JX459908) showed structural changes near and within epitope A (Figure 2). When the alanine insertion and SVTQTAT/A motif were introduced, several charged amino acids in GII.4 Sydney_2012 were replaced by neutral amino

acids (Figure 2). We also observed changes in the charge or hydrophobicity of amino acids in the monoclonal antibody binding epitope G (A356N) and within and around the HBGA binding regions D391N, S393D, and T395A, except in strains from South Africa (Appendix Figure 1). The alanine insertion in GII.4 San Francisco strains does not ablate binding to ligands found in porcine gastric mucin (Appendix Figure 2), which is consistent with ligand binding patterns known to correlate with susceptibility.

Conclusions

We report a novel norovirus GII.4 variant, named GII.4 San Francisco, detected in human stool specimens from patients with AGE on at least 3 continents during 2017–2022. The novel strains have a unique amino acid insertion in VP1 at the start of epitope A. We observed a similar unique insertion on epitope D in GII.4 variant Farmington Hills, which emerged in 2002, replacing the GII.4 US95–96 viruses, which had been circulating globally since 1995 (*13*). Whether the emerging GII.4 San Francisco strains will replace the current globally dominant GII.4 Sydney variant is not yet clear. Previous studies showed that epidemic GII.4 viruses diversified and spread over wide geographic areas for several years before epidemic emergence (*14*).

GII.4 viruses have always had strong immunodominance on epitope A, and alterations in epitope A residues has affected antibody responses (*15*). Addition of alanine at the start of epitope A and introduction of several neutral amino acids (SVTQTAT/A) before the insertion indicate major changes in the structure that could have an outsize effect on neutralizing antibody responses. GII.4 San Francisco strains showed mutations at residues S393D and T395A in epitope D. Those changes kept the ligand binding stabilizing function; epitope D also is a neutralizing epitope and an HBGA binding site (*4*). That finding further indicates that this virus has potential for increased spread and warrants additional antigenicity studies. Those data provide information for evaluation of norovirus vaccines that are currently in clinical trials.

In conclusion, the unique amino acid insertion in epitope A of VP1 together with a >5% aa difference from existing GII.4 variants confirmed that GII.4 San Francisco can be classified as a new GII.4 variant. This virus variant is circulating on at least 3 continents, North America, Europe, and Africa. Early detection and rapid assigning of an agreed upon name for future GII.4 variants will be crucial to assessing their pandemic potential.

Biographical Sketch

Dr. Chhabra is a microbiologist at the National Calicivirus Laboratory, Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention. Her primary research interest is molecular epidemiology of gastroenteritis viruses.

References

1. Chhabra P, de Graaf M, Parra GI, Chan MC, Green K, Martella V, et al. Updated classification of norovirus genogroups and genotypes. J Gen Virol. 2019;100:1393–406. https://doi.org/10.1099/jgv.0.001318
2. Farahmand M, Moghoofei M, Dorost A, Shoja Z, Ghorbani S, Kiani SJ, et al. Global prevalence and genotype distribution of norovirus infection in children with gastroenteritis: a meta-analysis on 6 years of research from 2015 to 2020. Rev Med Virol. 2022;32:e2237. https://doi.org/10.1002/rmv.2237
3. Parra GI, Tohma K, Ford-Siltz LA, Eguino P, Kendra JA, Pilewski KA, et.al. Minimal antigenic evolution after a decade of norovirus GII.4 Sydney_2012 circulation in humans. J Virol. 2023;97:e0171622. https://doi.org/10.1128/jvi.01716-22
4. Lindesmith LC, Boshier FAT, Brewer-Jensen PD, Roy S, Costantini V, Mallory ML, et al. Immune imprinting drives human norovirus potential for global spread. mBio. 2022;13:e0186122. https://doi.org/10.1128/mbio.01861-22
5. Cannon JL, Barclay L, Collins NR, Wikswo ME, Castro CJ, Magaña LC, et al. Genetic and epidemiologic trends of norovirus outbreaks in the United States from 2013 to 2016 demonstrated emergence of novel GII.4 recombinant viruses. J Clin Microbiol. 2017;55:2208–21. https://doi.org/10.1128/JCM.00455-17
6. Manouana GP, Nguema-Moure PA, Mbong Ngwese M, Bock CT, Kremsner PG, Borrmann S, et al. Genetic diversity of enteric viruses in children under five years old in Gabon. Viruses. 2021;13:545. https://doi.org/10.3390/v13040545
7. Brown JR, Roy S, Ruis C, Yara Romero E, Shah D, Williams R, et al. Norovirus whole-genome sequencing by SureSelect target enrichment: a robust and sensitive method. J Clin Microbiol. 2016;54:2530–7. https://doi.org/10.1128/JCM.01052-16
8. Mans J, Murray TY, Taylor MB. Novel norovirus recombinants detected in South Africa. Virol J. 2014;11:168. https://doi.org/10.1186/1743-422X-11-168
9. Parra GI, Squires RB, Karangwa CK, Johnson JA, Lepore CJ, Sosnovtsev SV, et al. Static and evolving norovirus genotypes: implications for epidemiology and immunity. PLoS Pathog. 2017;13:e1006136. https://doi.org/10.1371/journal.ppat.1006136

10. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol. 2018;35:1547–9. https://doi.org/10.1093/molbev/msy096

11. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 2021;30:70–82. https://doi.org/10.1002/pro.3943

12. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–9. https://doi.org/10.1038/s41586-021-03819-2

13. Widdowson MA, Cramer EH, Hadley L, Bresee JS, Beard RS, Bulens SN, et al. Outbreaks of acute gastroenteritis on cruise ships and on land: identification of a predominant circulating strain of norovirus—United States, 2002. J Infect Dis. 2004;190:27–36. https://doi.org/10.1086/420888

14. Ruis C, Lindesmith LC, Mallory ML, Brewer-Jensen PD, Bryant JM, Costantini V, et al. Preadaptation of pandemic GII.4 noroviruses in unsampled virus reservoirs years before emergence. Virus Evol. 2020;6:veaa067. https://doi.org/10.1093/ve/veaa067

15. Debbink K, Donaldson EF, Lindesmith LC, Baric RS. Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity. J Virol. 2012;86:1214–26. https://doi.org/10.1128/JVI.06189-11

Address for correspondence: Preeti Chhabra, Centers for Disease Control and Prevention, Mailstop H18-7, Atlanta, GA 30329-4018, USA; email: PChhabra@cdc.gov

October 2023

Congentital Syphilis

EMERGING INFECTIOUS DISEASES
Congenital Syphilis — October 2023

- Serotype Distribution and Disease Severity in Adults Hospitalized with *Streptococcus pneumoniae* Infection, Bristol and Bath, UK, 2006–2022

- Spike in Congenital Syphilis, Mississippi, USA, 2016–2022

- Carbapenem-Resistant *Klebsiella pneumoniae* in Large Public Acute-Care Healthcare System, New York, New York, USA, 2016–2022

- Posttransfusion Sepsis Attributable to Bacterial Contamination in Platelet Collection Set Manufacturing Facility, United States

- Effects of COVID-19 on Maternal and Neonatal Outcomes and Access to Antenatal and Postnatal Care, Malawi

- Emergence of SARS-CoV-2 Delta Variant and Effect of Nonpharmaceutical Interventions, British Columbia, Canada

- Community Outbreak of *Pseudomonas aeruginosa* Infections Associated with Contaminated Piercing Aftercare Solution, Australia, 2021

- Characteristics of and Deaths among 333 Persons with Tuberculosis and COVID-19 in Cross-Sectional Sample from 25 Jurisdictions, United States

- Sporadic Shiga Toxin–Producing *Escherichia coli*–Associated Pediatric Hemolytic Uremic Syndrome, France, 2012–2021
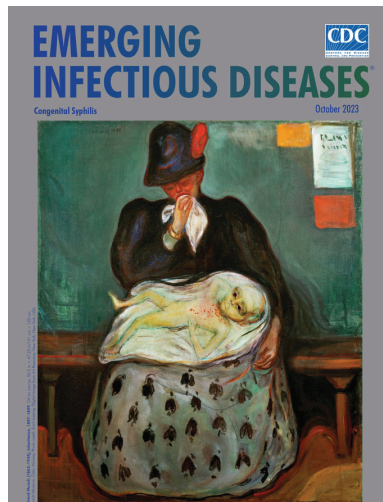
- Cycle Threshold Values as Indication of Increasing SARS-CoV-2 New Variants, England, 2020–2022

- Comprehensive Case–Control Study of Protective and Risk Factors for Buruli Ulcer, Southeastern Australia

- *Candida auris* Clinical Isolates Associated with Outbreak in Neonatal Unit of Tertiary Academic Hospital, South Africa

- Stability of Monkeypox Virus in Body Fluids and Wastewater

- Ancestral Origin and Dissemination Dynamics of Reemerging Toxigenic *Vibrio cholerae*, Haiti

- *Treponema pallidum* Detection at Asymptomatic Oral, Anal, and Vaginal Sites in Adults Reporting Sexual Contact with Persons with Syphilis

- Managing Risk for Congenital Syphilis, Perth, Western Australia, Australia

- Estimated Costs of 4-Month Pulmonary Tuberculosis Treatment Regimen, United States

- Human Tularemia Epididymo-Orchitis Caused by *Francisella tularensis* Subspecies *holartica*, Austria

- Imported Toxigenic *Corynebacterium Diphtheriae* in Refugees with Polymicrobial Skin Infections, Germany, 2022

- Expansion of Invasive Group A *Streptococcus* M1UK Lineage in Active Bacterial Core Surveillance, United States, 2019–2021

- Estimate of COVID-19 Deaths, China, December 2022–February 2023

- Mpox in Children and Adolescents during Multicountry Outbreak, 2022–2023

- Outbreak of Sexually Transmitted Nongroupable *Neisseria meningitidis*–Associated Urethritis, Vietnam

- *Pseudomonas aeruginosa* High-Risk Sequence Type 463 Co-Producing KPC-2 and AFM-1 Carbapenemases, China, 2020–2022

EMERGING INFECTIOUS DISEASES®

To revisit the October 2023 issue, go to:
https://wwwnc.cdc.gov/eid/articles/issue/29/10/table-of-contents