African horse sickness (AHS) is a hemorrhagic viral fever of horses. It is the only equine disease for which the World Organization for Animal Health has introduced specific guidelines for member countries seeking official recognition of disease-free status. Since 1997, South Africa has maintained an AHS controlled area; however, sporadic outbreaks of AHS have occurred in this area. We compared the whole genome sequences of 39 AHS viruses (AHSV) from field AHS cases to determine the source of 3 such outbreaks. Our analysis confirmed that individual outbreaks were caused by virulent revertants of AHSV type 1 live, attenuated vaccine (LAV) and reassortants with genome segments derived from AHSV types 1, 3, and 4 from a LAV used in South Africa. These findings show that despite effective protection of vaccinated horses, polyvalent LAV may, paradoxically, place susceptible horses at risk for AHS.
from the state veterinary service, and since March 2015, only during the period of low vector activity.

Since its creation in 1997, a total of 6 outbreaks of AHS in the AHS controlled area have been reported to OIE, in 1999, 2004, 2006, 2011, 2013, and 2014 (13–17). Before the 2014 outbreak, these outbreaks were assumed to be caused by illegal movement of viremic animals into the controlled area, although the source was established for only 2 of these outbreaks: a type 7 virus for the 1999 outbreak in the surveillance zone and a type 5 virus for the 2006 outbreak in the PZ (Figure 1) (18,19). Because the source of the viruses responsible for the other outbreaks was never established, the goal of our study was to further characterize the epidemiology of AHSV type 1 (AHSV-1) outbreaks in the controlled area by 1) whole-genome sequencing of viruses from individual outbreaks (2004, 2011, and 2014); 2) phylogenetic comparison of these sequences with those of the polyvalent AHSV-LAV and AHSV reference strains; 3) analysis of outbreak viruses for genome segment reassortment; 4) analysis of single-nucleotide variants (SNVs) associated with attenuation of AHSV-LAV to determine whether vaccine-derived viruses have reverted to virulence; 5) correlation of epidemiologic and clinical findings with molecular findings; and 6) confirmation of the source of the virus strains responsible for the 2004, 2011, and 2014 outbreaks of AHS in the controlled area.

### Materials and Methods

#### Virus Isolates

We sequenced complete genomes from 55 AHSV isolates collected during 1961–2014, including 39 field isolates of AHSV-1 from horses during the 2004 Stellenbosch (16 isolates), 2011 Mamre (7 isolates), 2014 Porterville (14 isolates), and 2014 Robertson (2 isolates) outbreaks of AHS in Western Cape Province of South Africa (Figure 1); AHSV LAV strains of types 1, 2, 3, 4, 6, 7, and 8; and Agricultural Research Council–Onderstepoort Veterinary Institute Laboratory reference strains for each of the 9 AHSV types. We included each of these virus isolates in the AHS genome sequencing Bioproject (http://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA271179) and identified each by a unique virus strain name (online Technical Appendix 1, http://wwwnc.cdc.gov/EID/article/22/12/16-0718-Techapp1.xlsx).
RNA Extraction, Identification, and Typing
We isolated individual viruses of each type included in the polyvalent AHSV-LAV independently, as previously described (20,21). We extracted genomic double-stranded RNA from all AHSV strains evaluated from virus-infected cells by using TRIzol reagent (Life Technologies, Johannesburg, South Africa). We identified and typed AHSV isolates by using group-specific (GS) real-time reverse transcription PCR (rRT-PCR) assays (22) and type-specific (TS) rRT-PCR assays targeting the gene encoding viral protein (VP) 2 (VP2) (23).

Genome Sequencing and Assembly
We prepared sequencing templates by using sequence-independent whole-genome RT-PCR amplification (24). We sequenced PCR amplicons on an Illumina MiSeq sequencer (Inqaba Biotechnical Industries, Pretoria, South Africa) by using the Nextera XT DNA sample preparation kit and 300-bp paired-end V3 Illumina chemistry. We analyzed Illumina sequence reads by using Genieous version 9 (http://www.geneious.com) (25). We used a combination of de novo assembly followed by mapping to obtain the full-length consensus genome sequences of each virus strain.

Phylogenetic Analysis
We aligned sequences of the concatenated whole virus genomes and individual genome segments by using MAFFT (http://mafft.cbrc.jp/alignment/software) (26) implemented within Geneious version 9 (25). We then used the Smart Model Selection program included in PhyML version 3 (http://www.atgc-montpellier.fr/phyml) (27) to identify the evolutionary models that best fit the individual sequence datasets by applying the corrected Akaike information criterion. We used the parameters from these models to construct maximum-likelihood trees by using PhyML version 3 (27) implemented within Geneious version 9 (25) with 1,000 bootstrap replicates to estimate branch support.

Genotype Group Analysis
We used RAMI (http://mbio-serv2.mbioekol.lu.se/rami.html) to align the concatenated whole genome sequence maximum-likelihood tree to genetically (and not evolutionarily) classify the sequences into genotype groups based on patristic distances (28). We ran RAMI with the patristic distance threshold set to 0.000459, enabling us to differentiate between genome sequences that differed from one another by as few as 16 nt variants.

Reassortment Analysis
We used Recombination Detection Program (RDP) version 4.63 (29) with default settings, except that we invoked the “scan for reassortment and recombination” setting to identify any reassortment between the gene segments of LAV strains of AHSV types 1, 3, and 4 and the 39 field isolate strains included in this study. We considered reassortment events detected by any of the 8 different recombination detection methods implemented in RDP (RDP, MAXCHI, and GENECONV methods in primary scanning mode and the BURT, Bootscan, CHIMAERA, SisScan, and 3SEQ methods in secondary scanning mode, each with a Bonferroni corrected p value cutoff of 0.05) to represent evidence of reassortment.

Nonsynonymous Single Nucleotide Variants
We aligned consensus concatenated whole virus genomes from 2 AHSV-1 laboratory strains (1/Lab/ZAF/62/OVI-HS29/62 and 1/Lab/ZAF/98/OBP-116) and the 39 field isolate strains by using MAFFT (26) and analyzed them by using the find variations/SNPs function in Geneious (25) with the find nonsynonymous polymorphisms only option enabled. We then compared the nonsynonymous SNVs in these sequences with the nonsynonymous SNVs previously associated with attenuation of AHSV-1 (24).

Results
Phylogenetic and Genotype Group Analysis
We used concatenated genome segments of 55 AHSV genomes to construct a maximum-likelihood phylogenetic tree incorporating best-fit substitution models (online Technical Appendix 2 Table 1, http://wwwnc.cdc.gov/EID/article/22/12/16-0718-Techapp2.pdf) to infer degrees of genetic relatedness (Figure 2). Genotype group analysis of patristic distances inferred from this maximum-likelihood tree by using RAMI (28) indicated that the AHSV strains isolated during the 2004, 2011, and 2014 outbreaks of AHS in the controlled area segregate into 3, 1, and 2 unique groups, respectively. Specifically, the groups were 1a, 1b, and 1c for the 2004 outbreak; 2 for the 2011 outbreak; 3a for the 2014 Porterville outbreak; and 3b for the 2014 Robertson outbreak (online Technical Appendix 2 Table 2). For the 2004 outbreak viruses, genotype group 1a includes 4 viruses that group closely with the AHSV-1-LAV strain, 1/Lab/ZAF/98/OBP-116; genotype group 1b includes 11 viruses that are also closely related to 1/Lab/ZAF/98/OBP-116; and genotype group 1c includes a single virus that segregates between 1/Lab/ZAF/98/OBP-116 and 4/Lab/ZAF/98/OBP-116. The Mamre outbreak viruses in genotype group 2 consist of 7 viruses that are all closely related to 1/Lab/ZAF/98/OBP-116. For the 2014 outbreak viruses, genotype group 3a includes 14 viruses that were all isolated from AHSV-infected horses in the Porterville area and genotype group 3b includes 2 viruses that were isolated from AHSV-infected horses in the Robertson area, with both groups of viruses being closely related to 1/Lab/ZAF/98/OBP-116.
the viruses included in genotype groups 1a, 2, 3a, and 3b also group with 1/Lab/ZAF/98/OBP-116, whereas those in genotype groups 1b and 1c group with the AHSV-3-LAV strain, 3/Lab/ZAF/98/OBP-116. For the gene encoding VP5, viruses included in all genotype groups except 1c group with 1/Lab/ZAF/98/OBP-116, whereas those in genotype group 1c group with the AHSV-4-LAV strain, 4/Lab/ZAF/98/OBP-116. For genes encoding NS3, viruses included in genotype groups 1a and 2 group with 1/Lab/ZAF/98/OBP-116, whereas those included in the remaining genotype groups group with 4/Lab/ZAF/98/OBP-116. Collectively, these data confirm that all 10 gene segments of
the viruses included in genotype groups 1a and 2 are probably derived from a most recent common ancestor closely resembling 1/Lab/ZAF/98/OBP-116; the viruses included in genotype groups 1b and 1c are probably reassortants derived from parental viruses very closely resembling 1/Lab/ ZAF/98/OBP-116, 3/Lab/ZAF/98/OBP-116, and 4/Lab/ ZAF/98/OBP-116; and viruses in genotype groups 3a and 3b are probably reassortants derived from parental viruses very closely resembling 1/Lab/ZAF/98/OBP-116 and 4/ Lab/ZAF/98/OBP-116 (Figure 3).

Figure 2. Whole-genome phylogeny of African horse sickness (AHS) viruses identified in AHS outbreaks in Western Cape Province, South Africa, 2004–2014. Maximum-likelihood phylogenetic tree indicating the genetic relationships of concatenated whole genome nucleotide sequences of AHS viruses from affected horses in the 2004, 2011, and 2014 outbreaks in the AHS controlled area in Western Cape Province to the AHS live, attenuated vaccine viruses and reference viruses. Green indicates vaccine-derived 1/Lab/ZAF/98/ OBP-116, blue indicates vaccine-derived 3/Lab/ZAF/98/OBP-116, and red indicates vaccine-derived 4/Lab/ZAF/98/OBP-116. Branches are scaled to represent numbers of inferred nucleotide differences per site. Branches supported by full maximum-likelihood bootstrap values >70% are indicated. Genotype groups are indicated at right. Scale bar indicates genetic distance.
Explicitly testing for intrasegment recombination and reassortment by using RDP4.63 (29) yielded no evidence of intracomponent recombination in any virus but strong evidence of reassortment in genotype group 1b, 1c, 3a, and 3b viruses (online Technical Appendix 2 Table 3). Genotype group 1b viruses have 6 genome segments (encoding VP2, VP3, VP5, VP6, NS1, and NS2) derived from a virus resembling 1/Lab/ZAF/98/OBP-116; 2 segments (encoding VP1 and VP7) derived from a virus resembling 3/Lab/ZAF/98/OBP-116 (multiple testing corrected p = 2.27 \times 10^{-12} and 1.13 \times 10^{-31}, respectively); 1 segment (encoding NS3) derived from a virus resembling 4/Lab/ZAF/98/OBP-116 (p = 9.31 \times 10^{-20}); and 1 segment (encoding VP4) that could plausibly have been derived from either 3/Lab/ZAF/98/OBP-116 or 1/Lab/ZAF/98/OBP-116 (p = 7.66 \times 10^{-4}) (Figure 4). Genotype group 1c viruses display a reassortant pattern resembling that of group 1b viruses except that the segment encoding VP5 is apparently derived from a virus resembling 5/Lab/ZAF/98/OBP-116 (p = 1.96 \times 10^{-26}). Genotype groups 3a and 3b viruses have 9 segments derived from a virus resembling 1/Lab/ZAF/98/OBP-116 and a single segment (NS3) derived from a virus resembling 4/Lab/ZAF/98/OBP-116 (p = 9.31 \times 10^{-20}) (Figure 4).

Several SNVs relative to the AHSV-LAV–derived viruses 1/Lab/ZAF/98/OBP-116 and 3/Lab/ZAF/98/OBP-116 are present in the NS1-encoding genes of viruses included in genotype groups 3a (2014 Porterville) and 3b (2014 Robertson) (online Technical Appendix 2 Table 4). Only a single nonsynonymous SNV exists between the NS1-encoding genes of 1/Lab/ZAF/98/OBP-116 and 3/Lab/ZAF/98/OBP-116 (NS1 I264T). All viruses included in genotype group 3a (2014 Porterville) have the I amino acid variant that is present in 1/Lab/ZAF/98/OBP-116, whereas viruses in genotype group 3b (2014 Robertson) include the T amino acid variant present in 3/Lab/ZAF/98/OBP-116. Viruses in the 3b genotype group (2014 Robertson) include \geq 2 synonymous SNVs and \geq 1 nonsynonymous SNV relative to 3/Lab/ZAF/98/OBP-116, which suggests that the NS1 gene of the virus strains in genotype group 3b are most probably derived from 3/Lab/ZAF/98/OBP-116, whereas those in genotype group 3a are more probably derived from 1/Lab/ZAF/98/OBP-116.

Seven nonsynonymous SNVs were identified between the whole genome sequences of the AHSV-1-LAV–derived virus, 1/Lab/ZAF/98/OBP-116, and its parental virus, 1/E.cab-tc/ZAF/62/OVI-HS29/62 (Table 1). SNVs are present at 4 of these 7 sites in the 4 viruses included in genotype group 1a. Intriguingly, 3 of these 4 changes are apparently reversions to the nonsynonymous SNV that is present in the virulent parental virus (I434T in VP5 and V81A and Q169R in VP6) and are therefore potentially reversion-to-virulence mutations. The 1 other SNV in the genotype group 1a viruses is site 201 in NS3, whereas in 1/Lab/ZAF/98/OBP-116 and 1/E.cab-tc/ZAF/62/OVI-HS29/62, a K and an M, respectively, are at this site, and in the group 1a viruses, an E is at this site. In 10 of the 11 field viruses in genotype group 1b, nonsynonymous SNVs were also detected at 4 of the 7 sites that differentiate the attenuated 1/Lab/ZAF/98/OBP-116 virus from its virulent parent, 1/E.cab-tc/ZAF/62/OVI-HS29/62. The remaining field virus in genotype group 1b, 1/E.cab-tc/ZAF/04/Vdm-E040065, includes 3 of these 4 SNVs. The I434T SNV in VP5 and the V81A and Q169R SNVs in VP6 of viruses in genotype group 1a are the same as those found in the genotype group 1b, 2, 3a, and 3b viruses. The K357N SNV in VP2 was detected only among viruses in genotype groups 1b and 1c. All the viruses included in genotype groups 1b, 1c, 3a, and 3b are also reassortants with an NS3-encoding segment derived from a virus resembling 4/Lab/ZAF/98/OBP-116; therefore, SNVs in this component of these viruses were not considered as genuine mutationally derived SNVs.

The 1/E.cab-tc/ZAF/04/Avt-E040061 strain in genotype group 1c has nonsynonymous SNVs at 3 of the 7 loci (K357N in VP2 and V81A and Q169R in VP6) but a
VP5-encoding gene apparently derived by reassortment from a virus resembling 4/Lab/ZAF/98/OBP-116, such that the SNVs in the VP5 of this strain were also not considered to be mutationally derived.

The 7 viruses included in genotype group 2 and the 16 viruses included in genotype groups 3a and 3b all exhibit potential reversion-to-virulence mutations at 3 of the 7 nonsynonymous SNV sites that differentiate the AHSV-1-LAV virus from its virulent parent (I434T in VP5 and V81A and Q169R in VP6). Additionally, a fourth SNV (K201N in NS3) at 1 of the 7 sites differentiating the AHSV-1 LAV from its parent (which had a K and an M, respectively, at this site) is also present in the genotype group 2 viruses.

**Quantification of the Outbreaks**

The epidemiologic parameters of the AHS outbreaks in the controlled area in 2004, 2011, and 2014 were inferred by using the current OIE case definition for AHS (11) (Table 2). Although the case-fatality rates (CFRs) were very high for the 2004 Stellenbosch (78.3%) and 2011 Mamre (76.2%) outbreaks, they were considerably lower for the 2014 Porterville (14.6%) and Robertson (4.5%) outbreaks. Additionally, the 2011 Mamre and 2004 Stellenbosch outbreaks were associated with the lowest vaccination rates among AHSV-infected horses (2.7% and 8.7%, respectively). Differences in the genetic constitution of the individual outbreak viruses could have been associated with the vastly different CFRs in each outbreak; however, whether these differences in CFRs are a consequence of lower virulence among the outbreak viruses or the result of existing vaccine-induced immunity in the exposed horses is unknown. Similarly, changes in the AHS case definition that only came into effect in 2008 (after the 2004 Stellenbosch outbreak) probably resulted in an underestimation of subclinical AHSV infections during that outbreak. Whereas during the Stellenbosch 2004 outbreak only clinically affected, deceased horses were classified
as having confirmed cases (13), major advances in AHS diagnostic testing (e.g., rRT-PCR–based methods) have occurred during the past 10 years that likely substantially increased the detection of subclinical infections by the time of the 2014 outbreaks (15,32).

**Discussion**

Whole-genome sequences were compared from 55 field, LAV, and laboratory strains of AHSV. The field viruses were obtained from horses during outbreaks of AHS of different clinical severity (CFRs ranging from 4.5% to 78.3%) in the AHS controlled area of South Africa during 2004, 2011, and 2014. Phylogenetic analyses confirmed that genetically distinct viruses were responsible for each outbreak and that these were all closely related to viruses contained in the trivalent AHSV-LAV (combination 1) used in South Africa. Evaluation of nonsynonymous SNVs confirmed some outbreak viruses to be revertants of the vaccine AHSV-1 strain toward the virulent parental type. Furthermore, some outbreak viruses were clearly reassortants with individual genome segments derived from multiple different virus types that are present in the trivalent vaccine preparation.

Potgieter et al. (24) hypothesized that changes in VP2 and VP5 can confer virulence or attenuation of individual AHSV strains, based upon comparisons of the consensus sequences of the genome of an attenuated AHSV-1 isolate (GenBank accession nos. FJ183364–FJ183373) and its virulent parent. Potgieter et al. (24) also proposed that virulence is related to tissue tropism because the outer capsid proteins are involved in cell entry and trigger apoptosis of host cells. Additionally, other studies have implicated NS3 as a determinant of AHSV virulence (33). The results of our study further confirm that changes in multiple VPs can affect the virulence of AHSV. Both reversion (to the virulent parental type) and novel SNVs were present in field-isolated viruses at various residue sites in VP2 (K357N in genotype group 1b viruses), VP5 (I434T in all field viruses evaluated except sample 1/E.cab-tc/ZAF/04/Avt- E040061), and VP6 (V81A in all field viruses and Q169R in all field viruses except sample 1/E.cab-tc/ZAF/04/Vdm-E040065) that differentiate the attenuated AHSV-1-LAV strain from its virulent parental strain. Furthermore, SNVs present at a site in NS3 (K201E in genotype group 1a viruses and K201N in genotype group 2) are potentially associated with reversion to virulence because of the effect of NS3 on virus release, membrane permeability, and viral yield (34). However, the determinants of AHSV virulence are probably complex and multigenic (24,34), which is consistent with the remarkable difference in CFRs between horses in the various outbreaks.

Given the genetic diversity of field strains of AHSV (14,24,35), our analyses overwhelmingly support the
premise that the potential reversion-to-virulence mutants and reassortants that we detected arose from viruses within the polyvalent AHSV-LAV formulation, and predominantly from AHSV-1-LAV. Although these mutants and reassortants most likely arose within vaccinated horses, the reason for the predominance of AHSV-1-LAV components in the emergent outbreak viruses is unknown. The data presented here also indicate that distinct founder events led to the expansion in Stellenbosch (2004) of viruses included in genotype groups 1a and 1b and, similarly, that the outbreaks in 2014 in Porterville (genotype group 3a) and Robertson (genotype group 3b) also probably originated independently from the LAV and were not from the spread of the same outbreak virus.

In summary, results of this study highlight the importance of genetic characterization of circulating strains of AHSV in epidemiologic investigations of AHS outbreaks. Although, the prevailing opinion in South Africa was that illegal movement of viremic equids into the AHS controlled area was responsible for the repeated occurrences of AHS in the controlled area, this is clearly not the only cause. Our data confirm that use of polyvalent AHSV-LAV can result in the emergence and spread of virulent viruses to adjacent susceptible horses, presumably by Culicoides midge vectors that are already resident within the AHS controlled area (36). Collectively, these findings have major implications for strategies to control AHS, both in AHS-endemic regions and during future incursions into currently AHSV-free areas. However, AHSV-LAV confers critical and effective protection for susceptible horses in AHS-endemic areas and, although potentially safer recombinant AHSV vaccines have proven effective in laboratory studies (37,38), these are not available commercially and they are yet to be evaluated in the field. Until alternative vaccines become commercially available, control of AHS will remain reliant on the use of AHSV-LAV coupled with the adoption of strategies to minimize the likelihood of natural dissemination of revertant and reassortant vaccine-derived viruses.

Acknowledgments
We thank the private-sector veterinarians, state veterinarians, and horse owners who collected samples and provided data used in this study.

Dr. Weyer is a veterinary research officer employed by the Equine Research Centre, Faculty of Veterinary Science, University of Pretoria, South Africa. She is authorized by the Western Cape provincial veterinary services to assist with equine movement control and disease surveillance within the province. Her areas of interest include African horse sickness epidemiology and other equine diseases, particularly those affecting the movement and trade of equines.

References


Address for correspondence: Alan J. Guthrie, University of Pretoria, Equine Research Centre, Faculty of Veterinary Science, Pretoria 0002, South Africa; email: alan.guthrie@up.ac.za

The Public Health Image Library (PHIL)
The Public Health Image Library (PHIL), Centers for Disease Control and Prevention, contains thousands of public health-related images, including high-resolution (print quality) photographs, illustrations, and videos. PHIL collections illustrate current events and articles, supply visual content for health promotion brochures, document the effects of disease, and enhance instructional media.

PHIL images, accessible to PC and Macintosh users, are in the public domain and available without charge.

Visit PHIL at http://phil.cdc.gov/phil

2096
Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 22, No. 12, December 2016