

https://dspace.org/

Epidemiology and genotypes of Enteroviruses in patients with respiratory illness in South Africa

Hellferscee, Orienka

2020
https://hdl.handle.net/10539/30452

Downloaded from a DSpace repository. https://dspace.org/

Epidemiology and genotypes of Enteroviruses in patients with respiratory illness in South Africa

Orienka Hellferscee

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirement for the degree of Doctor of Philosophy

Johannesburg, 2019

Declaration

I, Orienka Hellferscee, declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

6 December 2019, Johannesburg

International and national presentations from this study

1. **Orienka Hellferscee**, Florette Treurnicht, Stefano Tempia, Ebrahim Variava, Halima Dawood, Kathy Kahn, Shabir Madhi, Adam Cohen, Cheryl Cohen and Marietjie Venter. Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients Hospitalised with Severe Acute Respiratory Illness (Oral presentation)
    - 1st International Respiratory Pathogens Conference, Singapore 1-5 September 2015

2. **Orienka Hellferscee**, Stefano Tempia, Ebrahim Variava, Halima Dawood, Shabir A. Madhi, Mignon du Plessis, Florette K. Treurnicht, Cheryl Cohen. Enterovirus Genotypes Detected Among Patients With Severe Acute Respiratory Illness, Influenza-like Illness And Asymptomatic Individuals In South Africa, 2012-2014 (Poster presentation)
    - 5th African Network for Influenza Surveillance and Epidemiology (ANISE) meeting, Kigali Rwanda, 8 - 11 March 2016
    - WITS research day 2016, Johannesburg, 1 September 2016

3. **Orienka Hellferscee,** Florette Treurnicht, Sibongile Walaza, Mignon du Plessis, Anne von Gottberg, Jocelyn Moyes, Halima Dawood, Ebrahim Variava, Marthi Pretorius, Marietjie Venter, Cheryl Cohen, Stefano Tempia. The attributable fraction of rhinovirus detection to mild and severe respiratory illness among HIV-infected and HIV-uninfected South African patients, 2013-2015" (Poster presentation)
    - 6th African Network for Influenza Surveillance and Epidemiology (ANISE) meeting, Antananarivo Madagascar, 19-21 March 2018
    - WITS research day 2018, Johannesburg, 6 September 2018

Publications directly contributing to this thesis and role of the student in these studies

Data from all participants are from respiratory illness surveillance conducted over a six-year period (2009-2015) for which I was responsible for performing and overseeing virology laboratory testing of surveillance samples from 2011 for the duration of the study.

In addition, I conducted all characterisation testing, phylogenetic analyses and statistical analyses for the publications included in this thesis and drafted and revised the published manuscripts.

1. **Hellferscee O**, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, Cohen AL, Pretorius M, Cohen C, Madhi SA, Venter M. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza and Other Respiratory Viruses. 2017 May; 11(3):211-219

2. **Hellferscee O**, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, Madhi SA, du Plessis M, Cohen C, Treurnicht FK. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. Journal of Medical Virology 2017 Oct; 89(10):1759-1767

3. **Hellferscee O**, Treurnicht FK, Walaza S, du Plessis M, von Gottberg A, Moyes J, Dawood H, Variava E, Pretorius M, Venter M, Cohen C, Tempia S. The fraction of rhinovirus detections attributable to mild and severe respiratory illness in a setting of high human immunodeficiency virus prevalence, South Africa, 2013–2015. Journal of Infectious Diseases 2019 May 5;219(11):1697-1704

4. **Hellferscee O**, Janse van Vuren P, du Plessis M, Cohen C, Treurnicht FK. Complete genome sequence analysis of human echovirus 30 isolated from patients with acute respiratory illness in South Africa, 2013. (In preparation)

Ethics clearance certificate

UNIVERSITY OF THE
WITWATERSRAND
JOHANNESBURG

R14/49 Ms O Hellferscee

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180386

NAME: (Principal Investigator)	Ms O Hellferscee
DEPARTMENT:	National Institute for Communicable Diseases Centre for Respiratory Diseases and Meningitis National Health Laboratory Service
PROJECT TITLE:	Epidemiology and genotypes of enteroviruses in patients with respiratory illness in South Africa
DATE CONSIDERED:	06/04/2018
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M140824 and antecedents
SUPERVISOR:	Drs C Cohen, M du Plessis and F Treurnicht
APPROVED BY:	Professor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	06/04/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

Abstract

Globally, lower respiratory tract infections (LRTI) are the leading cause of death in children aged less than 5 years. This study focuses on enteroviruses, which are a known cause of respiratory illness but whose epidemiology and molecular characterization have not been well described in South Africa.

The first aim of this study was to describe enterovirus genotypes circulating among South African patients hospitalized with severe respiratory illness (SRI). Upper respiratory samples were tested for the presence of enteroviruses using polymerase chain reaction (PCR) and characterized using a partial VP1 gene segment through Sanger sequencing. Enteroviruses were identified in 8.2% of patients with SRI. We demonstrated a high diversity, 17 genotypes, from all four enterovirus species circulating in South Africa.

The second aim was to use Sanger sequencing to characterize the enterovirus genotypes circulating among South African patients with influenza-like illness (ILI), SRI, and asymptomatic individuals. We found high genetic diversity (33 genotypes) but no association of enterovirus species with disease severity; however, some genotypes (echovirus 30 (E30), coxsackie virus B5, and EV-D68) were more prevalent in symptomatic cases. The E30 strains formed a new cluster proposed to be named genotype k.

The third aim described a novel approach for next generation sequencing of E30, based on sequence-independent single primer amplification (SISPA). Complete full genome phylogenetic analysis supports the designation of genotype k, and distinct grouping was confirmed that could be a result of the diversity of different subgenomic regions in E30 driven by different evolutionary paths. Five E30 whole genomes were submitted to GenBank, contributing to global data.

The fourth aim was to determine the fraction of rhinovirus (the most commonly identified species of Enterovirus) detection associated with mild or severe

respiratory illness. Rhinovirus was detected in 17.4% (368/2120), 26.8% (979/3654) and 23.0% (1003/4360) of controls, ILI and SRI, respectively. The overall RV-AF (ILI: 54.6%; severe acute respiratory illness (SARI): 54.3%) were similar to those reported in a previous study conducted in South Africa (ILI: 52.0%; SARI: 46.9%), but was lower compared to those of other important respiratory pathogens such as influenza viruses (ILI: 93.3%; SARI: 86.3%) and respiratory syncytial virus (ILI: 63.1%; SARI: 83.7%) (21). Even though the attributable fraction was lower than influenza and RSV, the attributable burden of rhinoviruses (RV-AF) was equivalent to influenza and RSV, because of the high rhinovirus detection rate.

In conclusion, we found (i) no difference in enterovirus genotypes detected in mild versus severe disease; (ii) that outbreak-associated EV strains like EV-D68 are circulating in South Africa; (iii) that E30 strains from respiratory illness cases in South Africa belong to a newly proposed genotype k; and (iv) that rhinovirus species contribute to mild and severe respiratory illness.

Acknowledgements

Alle eer aan my hemelse Pappa vir krag en verstand - “Ek is tot alles instaat deur U wat my krag gee”

My gesin – Alttden/Kayden/Laylah – dankie vir al die tyd wat julle my gegun het en vir al die koffie aandra. Girls, my “boek” is uiteindelik klaar en Lief, ek belowe, hierdie is die laaste ene ☺

My ouers – Pappie/Moemie, julle eindelose oproepe, ondersteuning en aanmoediging is van onskatbare waarde

My swes – Debs – dankie dat jy my nommer een cheerleader is!

Darinka/Jo-Andrie/Cardia – Buddies, dankie vir al die ondersteuning en aanmoediging, julle het ‘n spesiale plekkie in my hart

Amelia – my werk mamma – dankie vir al jou ondersteuning en dat jy soms meer werk moes aanvat sodat ek kon opskryf

My supervisors – I could not have asked for a better team:

- Cheryl/Florette/Mignon – thank you for believing in me as a scientist. Thank you for all your kind words and motivation and always making time to brainstorm, I am truly blessed to be mentored by the best
- Anne – thank you for all your guidance and teaching me how to think outside the box

Stefano/Meredith – thank you for all the reviews as part of writing group and Stefano thank you for teaching me statistics

Table of contents

Declaration……..... i
International and national presentations from this study……...……. ii
Publications directly contributing to this thesis and role of the student in these studies……..…….. iii
Ethics clearance certificate…….. v
Abstract…….. vi
Acknowledgements…….. viii
Table of contents……...……..... ix
List of figures……... xiv
List of tables…….. xx
List of abbreviations……... xxii

Chapter 1 Literature Review……... 1

1.1 Global overview of respiratory tract infection……..... 1
1.2 Respiratory tract infections in South Africa……... 2
1.3 Family *Picornaviridae*…… 3
1.3.1 Genus Enterovirus…… 5
1.3.2 Enterovirus species……... 6
1.3.2.1 Clinical manifestations and seasonality…… 6
1.3.2.2 Enterovirus D68…… 7
1.3.2.3 Echovirus 30…… 8
1.4 Rhinovirus species……..... 9
1.4.1 Clinical manifestations and seasonality…… 9
1.4.2 Does rhinovirus contribute to the burden of severe respiratory illness?..... 10
1.5 Viral genomics, viral evolution and recombination events ……... 11
1.6 Aims and objectives……. 13

Chapter 2 Materials and Methods…… 15

2.1 Description of the surveillance programmes and case definitions…… 15
2.1.1 Severe acute respiratory illness (SARI) surveillance…….. 15
2.1.2 Severe chronic respiratory illness (SCRI) surveillance…….. 16
2.1.3 Severe respiratory illness (SRI) surveillance…….. 16
2.1.4 Influenza-like illness (ILI) surveillance…….. 17
2.1.5 Asymptomatic individuals…….. 17
2.2 Study procedures……... 17
2.3 Description of sample collection……... 18
2.4 Laboratory methods and procedures……. 19
2.4.1 Total nucleic acid extraction……... 19
2.4.2 Laboratory methods for enterovirus and rhinovirus……...…… 20
2.4.3 Sanger sequencing of enterovirus VP1 gene region……... 21
2.4.3.1 Assembly and phylogenetic analysis of enterovirus partial VP1... 23
2.4.4 Enterovirus isolation……...… 23
2.4.5 Enterovirus genome sequencing……..… 23
2.4.5.1 Viral enrichment and sequence independent single primer amplification (RACE-SISPA) of echovirus 30……...… 23
2.4.5.1.1 Viral enrichment……. 23
2.4.5.1.2 Rapid amplification of cDNA (RACE)……...… 24
2.4.5.1.3 Sequence independent single primer amplification (SISPA)……...… 24
2.4.5.2 Next generation sequencing on the MiSeq platform…….. 26
2.4.5.3 Data analysis: Trimming, mapping of the reads and assembly….. 26
2.4.6 Phylogenetic analysis of echovirus 30 whole genome…… 27
2.5 Statistical data analysis……. 27

Chapter 3 Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients with Severe Acute Respiratory Illness, 2009-2011…… 30

3.1 Introduction……..... 30
3.2 Materials and Methods……..….... 31
3.2.1 Study design and population……...… 31

3.2.2 Sample selection and detection of enterovirus infection……. 31
3.2.3 Amplification and sequencing of enteroviruses partial VP1…..…..… 32
3.2.4 Statistical analysis……... 32
3.3 Results……..... 32
3.3.1 Clinical and epidemiological characteristics of patients hospitalised with enterovirus-associated SARI……...… 32
3.3.2 Molecular epidemiology of enterovirus species A to D…….. 33
3.3.3 Enterovirus serotypes and circulation of EV-D68 clades during 2009-2011……...… 36
3.4 Discussion……..… 41

Chapter 4 Enterovirus Genotypes among Patients with Severe Acute Respiratory Illness, Influenza-Like Illness and Asymptomatic Individuals in South Africa, 2012-2014……..… 43

4.1 Introduction……...… 43
4.2 Materials and Methods…….. 44
4.2.1 Study design and population……...… 44
4.2.2 Sample selection and detection of enterovirus infection……..… 44
4.2.3 Amplification and sequencing of enteroviruses…….. 44
4.2.4 Statistical analysis……...… 45
4.3 Results……... 45
4.3.1 Epidemiology of the study population……..... 45
4.3.2 Phylogenetic comparison of enterovirus genotypes identified during June 2012 – July 2014……. 46
4.3.3 Enterovirus group and genotypes…… 49
4.3.4 Molecular evolution of echovirus 30, coxsackie virus B5 and EV-D68 genotypes…….. 53
4.3.4.1 Echovirus 30……...… 53
4.3.4.2 Coxsackie virus B5……..… 53
4.3.4.3 EV-D68……...… 56
4.4 Discussion……..… 58

Chapter 5 Genome Analysis of Human Echovirus 30 Isolated from Patients with Acute Respiratory Illness in South Africa, 2013……..… 61

5.1 Introduction……... 61
5.2 Material and Methods..…… 63
5.2.1 Study design and population……..…….…... 63
5.2.2 Sample selection and detection of enterovirus infection…….….. 63
5.2.3 Amplification and sequencing of enteroviruses…… 63
5.2.4 Enterovirus isolation and genome sequencing…….. 63
5.3 Results…….. 64
5.3.1 Enterovirus isolation……..… 64
5.3.2 Characteristics of the echovirus 30 positive cases……. 64
5.3.3 Genome sequencing……... 65
5.3.4 Complete genome analysis of echovirus 30 samples…… 66
5.3.5 Phylogenetic analysis…… 68
5.3.5.1 Phylogenetic analysis of complete genomes…...……... 68
5.3.5.2 Phylogenetic analysis of sub-genomic P1 region……... 70
5.3.5.3 Phylogenetic analysis of sub-genomic P2 and P3 regions …… 71
5.3.5.4 VP1 sub-genomic analysis……...……... 73
5.3.6 Investigation of E30 genomes for recombination……..… 74
5.4 Discussion……..… 76

Chapter 6 The Fraction of Rhinovirus Detections Attributable To Mild and Severe Respiratory Illness in a Setting of High Human Immunodeficiency Virus Prevalence, South Africa, 2013–2015……... 79

6.1 Introduction……...… 79
6.2 Materials and Methods…….…. 80
6.2.1 Study design and population……..… 80
6.2.2 Sample selection and detection of rhinovirus infection…….. 80
6.2.3 Statistical analysis……..… 81
6.3 Results…….. 81
6.3.1 Study population…… 81

6.3.2 Rhinoviruses detection and attributable fraction……... 83
6.4 Discussion……...….. 91

Chapter 7 Conclusions…….…… 94

References……..... 97

List of figures

Figure	Title	Page
Figure 1.1	Negative-contrast electron micrograph of enterovirus particles; the bar represents 100 nm (courtesy of Ann C. Palmenberg)……	1
Figure 1.2	Diagrammatic presentation of the enterovirus capsid and genome structure (17)…….....	4
Figure 1.3	Pictures of the different human enterovirus species (family *Picornaviridae*, genus enterovirus). Species, type, virus name abbreviation and the virion's average diameter (Æ) are indicated (22)……..……..	6
Figure 2.1	Map of provinces in South Africa – indicating the two surveillance sites (courtesy: Sibongile Walaza, Centre for Respiratory Diseases and Meningitis, NICD)…….....	15
Figure 3.1	Molecular characterization of enterovirus clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implemented in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Black dots = South African samples 2009-2011……....	34
Figure 3.2	Epidemiologic graphs showing: (A) Distribution of enterovirus in South Africa during 2009-2011 and (B) Distribution of enterovirus species detected in South Africa during 2009-2011……	35
Figure 3.3	Molecular characterization of EV-D68 clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed	

using neighbor-joining methods as implicated in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Scale bar indicates number of nucleotide substitutions per site. Black dots = South African strains 2009-2011; Open dots = South African strains 2000-2001. Clades B1, B3, C A1 and partial B2 and A2 have been compressed to show comparison of South African strains to strains from other countries (country and year in blocks). Total number of sequences = 312 (representative strains from 2000-2016)……..... 40

Figure 4.1 Monthly number and overall detection rate of enterovirus (EV)-positive cases among controls (A), influenza-like illness (B) and severe acute respiratory patients (C), Klerksdorp and Pietermaritzburg, South Africa, 2012-2014…… 48

Figure 4.2 Phylogenetic analysis of enterovirus genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denotes serotypes identified in SARI patients, those with grey closed circles denotes genotypes identified in ILI patients while those with open circles denoted genotypes identifies in asymptomatic individuals. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree……... 50

Figure 4.3 The genotypes and species (EV-A to EV-D) of enteroviruses identified in patients with severe acute respiratory illness (SARI), influenza-like illness (ILI) and asymptomatic individuals (C) from June 2012 to July 2014, South Africa…….. 52

Figure 4.4 Distribution of enterovirus (EV) species (A-D) in South Africa during June 2012 to July 2014; dnt=did not genotype……..... 52

Figure 4.5 Phylogenetic analysis of echovirus 30 genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site……..... 54

Figure 4.6 Phylogenetic analysis of coxsackie virus B5 genogroups by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site…… 55

Figure 4.7 Phylogenetic analysis of EV-D68 lineages by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Sequences with open circles are EV-D68 strains previously identified from South Africa. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site……... 57

Figure 5.1 Genome structure of echovirus 30 (permission from ViralZone) (VPg – Viral protein genome linked; IRES – internal ribosomal entry site; RdRp – RNA dependent RNA polymerase)…… 62

Figure 5.2 Electrophoresis run from the Bioanalyzer indicating smear between 300-1000bp……... 65

Figure 5.3 Schematic representation of genome coverage after reference mapping against the E30 prototype, Bastianni strain (AF311938)……... 67

Figure 5.4 Entropy plot indicating the difference of the South African echovirus 30 strains across the polyprotein……. 66

Figure 5.5 Genomic (6580 bp) relationship of 5 South African echovirus 30 (E30) strains compared to 19 other E30 strains available in GenBank database. Genome sequences were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstraps replicated). The scale bar represents the genetic distance. Bootstrap values >65% are shown. ● South African strains from this study……. 69

Figure 5.6 Phylogenetic relationships based on the complete P1 sub genomic region (2580 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study…….. 70

Figure 5.7 Phylogenetic relationships based on the complete P2 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study……. 71

Figure 5.8 Phylogenetic relationships based on the complete P3 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study…….. 72

Figure 5.9 Phylogenetic relationships based on the complete VP1 genomic region (876 bp) of echovirus 30 (E30) strains. Five South African and 58 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis. Bold line indicates the proposed genotype k……..... 73

Figure 5.10 Topological incongruence report. LHS – left-hand side; RHS – right-hand side……. 74

Figure 5.11 Phylogenetic relationships based on 4 partitions generated using the significant breakpoints across the echovirus 30 (E30) genome. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis……..…….. 75

Figure 6.1 Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among outpatients with influenza-like illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected patients……. 88

Figure 6.2 Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among patients hospitalised with severe respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected……..... 89

List of tables

Table	Title	Page
Table 3.1	Clinical and epidemiological characteristics of patients hospitalised with enterovirus associated severe acute respiratory illness at surveillance sites, South Africa, 2009-2011……..	37
Table 3.2	Demographic and clinical characteristics of ten patients in whom EV-D68 was identified in South Africa…….....	39
Table 4.1	Association of enterovirus species (EV-A to EV-D) with influenza-like illness (ILI) and severe acute respiratory illness (SARI) compared with asymptomatic controls, South Africa, 2009-2010 (bold indicate factors significant at p<0.05)……	51
Table 5.1	Concentration of double stranded cDNA generated using RACE-SISPA comparing enriched and non-enriched viral samples as well as total reads generated with MiSeq……..	64
Table 6.1	HIV prevalence among outpatients with influenza-like illness and inpatients with severe (acute or chronic) respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013-December 2015……..	81
Table 6.2	Asymptomatic individuals positive for rhinovirus infection, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……...	83
Table 6.3	Rhinoviruses detection and rhinoviruses attributable fraction among outpatients with influenza-like illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……	84

Table 6.4 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015…….. 85

Table 6.5 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe acute respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……... 86

Table 6.6 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe chronic respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015…… 87

List of abbreviations

%	percentage
<	less than
≥	greater than or equal to
°C	degree Celsius
µg/ml	microgram per microliter
µl	microliter
Æ	average diameter
AF	attributable fraction
Ag/Ab	antigen and antibody
BLAST	Basic Local Alignment Search Tool
bp	base pairs
CD4+	T-cell CD4 T-helper cells
cDNA	complementary DNA
CI	confidence interval
Ct	cycle threshold
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E30	Echovirus 30
e.g	for example
EIA	enzyme immuno assay
ELISA	enzyme-linked immunosorbent assay
EV-A	enterovirus species A
EV-B	enterovirus species B
EV-C	enterovirus species C
EV-D	enterovirus species D
EV-D68	enterovirus D68
g	relative centrifugal force
GARD	Genetic Algorithm for Recombination Detection
HEp-2	Human epithelial type 2
HIV	human immunodeficiency virus
HPD	high probability density
i.e	in other words
ICU	intensive care unit
ILI	influenza-like illness
IRES	Type I internal ribosomal entry site
LRTI	lower respiratory tract infection
mM	millimolar
mRNA	messenger ribonucleic acid

NGS	next generation sequencing
NHLS	National Health Laboratory Service
NICD	National Institute for Communicable Diseases
nm	Nanometer
nt	nucleotides
NP/OP	combined nasopharyngeal and oropharyngeal specimen
OR	odds ratio
ORF	open reading frame
PCR	polymerase chain reaction
pmol	picomole
qRT-PCR	real-time reverse transcription polymerase chain reaction
RACE-SISPA	rapid amplification of cDNA end sequence independent single primer amplification
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
RV-A	rhinovirus species A
RV-AF	rhinovirus attributable fraction
RV-B	rhinovirus species B
RV-C	rhinovirus species C
SARI	severe acute respiratory illness
SCRI	severe chronic respiratory illness
SDS	sodium dodecyl sulfate
SISPA	sequence independent single primer amplification
SRI	severe respiratory illness
TNA	total nucleic acid
U	unit
UNICEF	United Nations Children's Fund
USA	United States of America
UTR	untranslated region
V	volts
VPg	genomic Viral Protein
vs.	versus
WHO	World Health Organization
µM	micromolar

Chapter 1

Literature Review

1.1 Global overview of respiratory tract infection

Lower respiratory tract infections (LRTI) are inflammation of the airways or pulmonary tissue, resulting from an infection below the level of the larynx and include conditions like bronchiolitis, bronchitis, laryngotracheobronchitis (croup) and pneumonia (4). LRTIs are a substantial public health problem and a leading cause of illness and death in people of all ages. LRTI's cause more than 2.75 million deaths worldwide, making this the 5th leading cause of death overall and the leading infectious cause of death in children less than 5 years of age (5). Despite substantial reductions in LRTI mortality in many countries, the burden remains high, particularly in areas of low sociodemographic development and has increased in some populations, particularly the extreme of ages – very young and elderly people (5).

Estimates of the global burden of LRTI's will be improved by more high-quality data on mortality, morbidity and aetiologies, especially in sub-Saharan Africa where the burden is highest and data are scarce. Improvements in diagnostics for LRTI's aetiologies, including those appropriate for children younger than five years of age to better understand the contribution of each aetiology to the LRTI burden, will help guide targeted interventions such as vaccination (5).

1.2 Respiratory tract infections in South Africa

LRTI is the most common cause of hospitalisation in human immunodeficiency virus (HIV)-infected children in sub-Saharan Africa and HIV infection is associated with increased severity of LRTI and higher case fatality ratios (6). The relative contribution of viral and bacterial aetiologies to the syndrome of pneumonia also varies by HIV status. Data from a severe acute respiratory illness (SARI) surveillance system, a prospective hospital based sentinel surveillance system, estimates the overall incidence of LRTI in South African children less than five years of age to be 2530-3173 per 100,000 with the highest incidence in children aged less than one year of age (annual range 8446-10532 per 100,000). The risk of hospitalisation with LRTI was up to three times higher in HIV-infected compared to HIV-uninfected children (7).

Influenza-like illness (ILI) in outpatients causes large numbers of illness episodes in South Africa and is a major reason for health facility usage in low and middle-income countries, contributing to the public health burden in terms of absenteeism from work, health-care and economic cost. Respiratory viruses associated with acute LRTI and ILI include influenza virus A and B, respiratory syncytial virus, parainfluenza virus types 1, 2 and 3, adenovirus, human metapneumovirus, enterovirus and rhinovirus (8,9).

Two of these viruses, influenza virus and respiratory syncytial virus, have been well characterized as causes of pneumonia in South Africa (10,11) and data from these studies have been used to guide vaccine policy and candidate vaccine strain selection (12). This fits well with the World Health Organization (WHO) and United Nations Children's Fund (UNICEF) integrated global action plan for pneumonia which aims to accelerate pneumonia control with a combination of interventions to protect, prevent and treat pneumonia in children (13). However, a number of other viruses has been suggested to cause pneumonia or ILI and have candidate vaccines in development, but these are not well described in South Africa. This includes viruses in the family *Picornaviridae*, genus Enterovirus – enteroviruses and rhinoviruses.

1.3 Family *Picornaviridae*

The family *Picornaviridae* consists of small, icosahedral non-enveloped viruses, 30-32 nm (14) (Figure 1.1). Picorna is derived from the prefix “pico” which designates a very small unit of measurement (10^{-12}) and RNA – therefore directly translated meaning very small RNA. The genome consists of a single-stranded non-segmented positive sense RNA of approximately 7500 nucleotides, containing of a single open reading frame (ORF). The ORF encoding a polyprotein flanked on either side by untranslated regions (UTRs) which contains regions of secondary structure, which are essential to genome function (14).

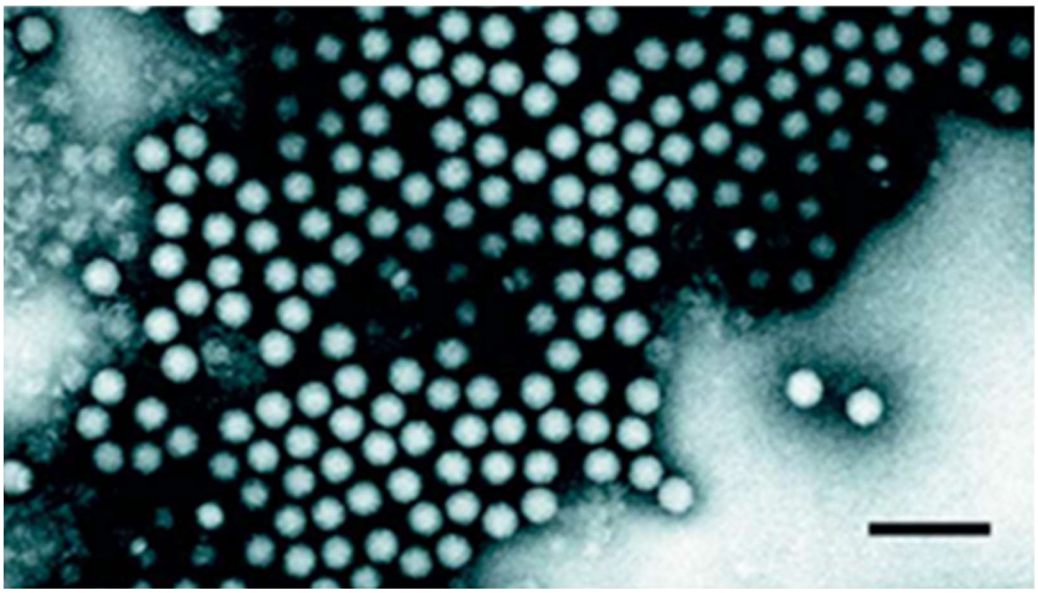


Figure 1.1: Negative-contrast electron micrograph of enterovirus particles; the bar represents 100 nm (courtesy of Ann C. Palmenberg).

The polyprotein is initially processed into three precursor proteins: P1, P2 and P3. The P1 genomic region encodes four capsid or structural proteins (VP1, VP2, VP3 and VP4) while the P2 and P3 genomic regions encode seven nonstructural proteins (2A, 2B, 2C, 3A, 3B, 3C and 3D) associated with replication (2,14,15). The capsid is made up of 60 copies of each of the four capsid proteins assembled into the icosahedral virion structure, and only the VP1, VP2 and VP3 proteins are exposed on the capsid surface (15) (Figure 1.2).

The viral genomic RNA has a viral protein (VPg) at its 5` end instead of a -methylated nucleotide cap structure. The long 5' UTR contains a type I internal ribosomal entry site (IRES), which allows direct translation of the polyprotein. The ORF also contains a poly(A) tail located after the 3' terminal sequence (2,14). The virion RNA is infectious, serving as both the genome and the viral mRNA (16).

The orthologous (genes in different species that evolved from a common ancestral gene by speciation) proteins VP2, VP3, VP1, 2C, 3C and 3D are conserved for the *Picornaviridae* family and may be used for sequence alignments between viruses, whereas VP4, 2A, 2B, 3A and 3B are highly divergent among the members of different picornavirus genera. None of the viral proteins are glycosylated (14). The family comprises 35 genera and human enteroviruses and rhinoviruses are members of the genus Enterovirus in the family *Picornaviridae* (14).

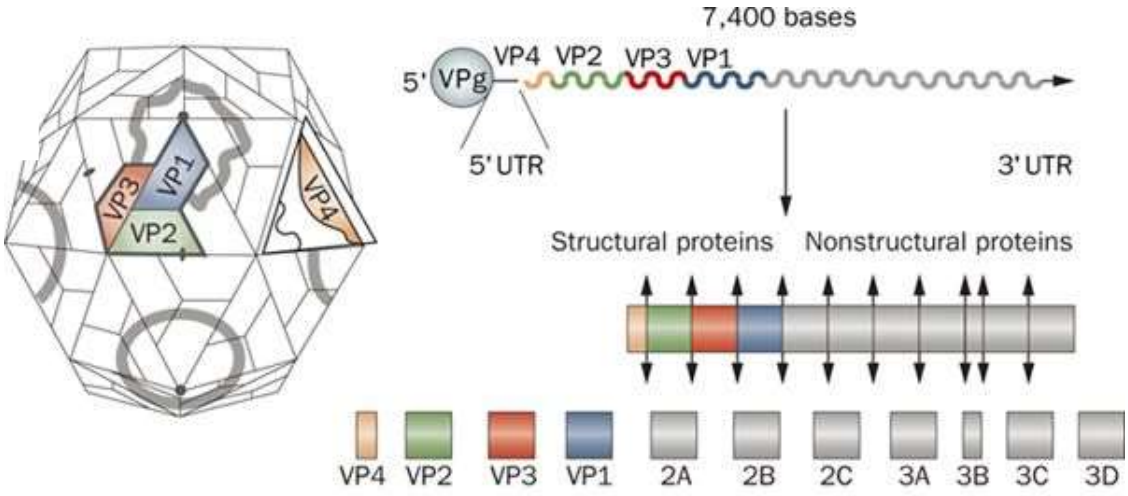


Figure 1.2: Diagrammatic presentation of the enterovirus capsid and genome structure (17).

1.3.1 Genus Enterovirus

Entero comes from the Greek word “enteron” which means intestine, which is an appropriate name since enteroviruses multiply primarily in the gastrointestinal tract or the upper respiratory tract or sometimes both, but they can also multiply in other tissues, e.g. nerve and muscle. Infections with enteroviruses are generally cytolytic, but persistent infections are common (14). Clinical manifestations include the common cold, mild meningitis, encephalitis, myelitis, myocarditis and conjunctivitis although some infections may frequently be asymptomatic (14). Transmission is horizontal, mainly faecal-oral, fomite or airborne routes (14).

Because of the high sequence diversity within the VP1 capsid region and genotype specific neutralization profiles, more than 300 virus types have been characterized in the genus Enterovirus (2). The genus genetically consists of about 75 enterovirus genotypes and 100 rhinovirus genotypes that cluster into 7 species (4 enterovirus species namely enterovirus A [EV-A], enterovirus B [EV-B], enterovirus C [EV-C], enterovirus D [EV-D] and 3 rhinovirus species namely rhinovirus A [RV-A], rhinovirus B [RV-B] and rhinovirus C [RV-C]) that infect humans (14) (Figure 1.3). Other types and genotypes include enterovirus E: 5 types, enterovirus F: 7 types, enterovirus G: 20 types, enterovirus H: 1 type, enterovirus I: 1 type, enterovirus J: 6 types, and at least 3 types that await classification. Four enterovirus species (A to D) are human pathogens but other enterovirus species (G to L) have been reported from animals (i.e. swine, monkey and rodent) (14,18–20).

The 5′-UTR of human rhinoviruses is shorter than that of enteroviruses, due to a deletion of approximately 100 nucleotides (nt) between the IRES and the translation start site (14). Human enteroviruses and rhinoviruses share identical genomic organization, have similar functional RNA secondary structures and are classified in the same genus because of their high sequence homology (14). However, despite their common genomic features, these two groups of viruses have different phenotypic characteristics including acid stability: virions of most

enterovirus species are stable at pH 3.0, while those of the rhinovirus species are unstable below pH 5-6 (14).

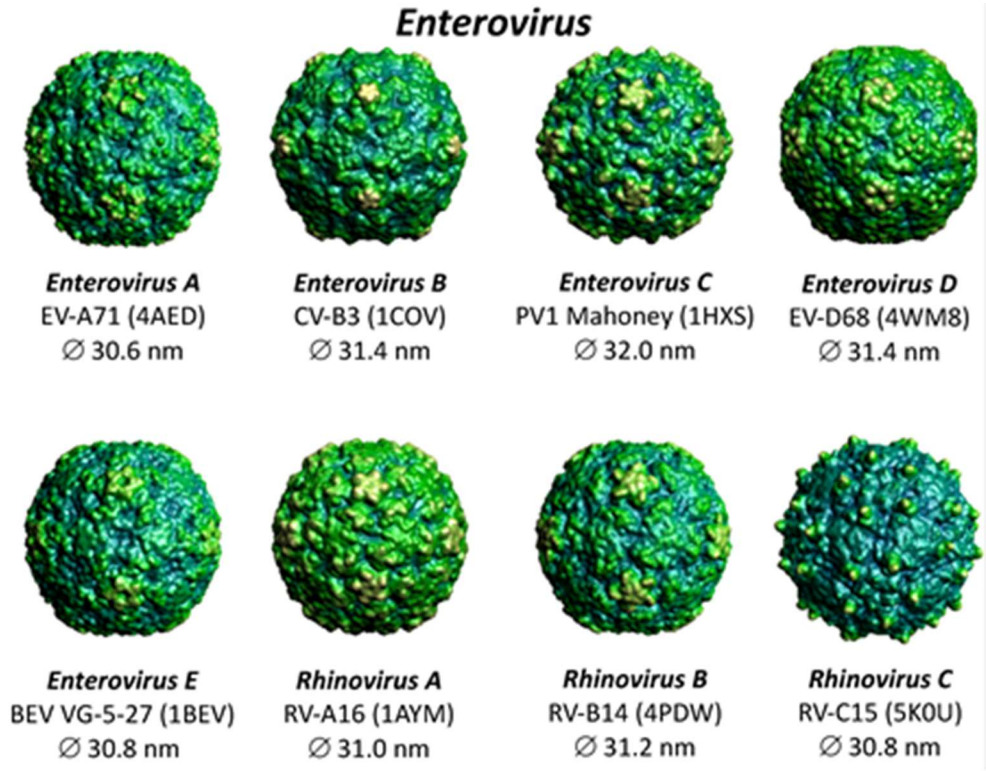


Figure 1.3: Pictures of the different human enterovirus species (family *Picornaviridae*, genus enterovirus). Species, type, virus name abbreviation and the virion's average diameter (Æ) are indicated (22).

1.3.2 Enterovirus species

1.3.2.1 Clinical manifestations and seasonality

Most human enteroviruses enter the body via the alimentary or respiratory tract, and spread by both faecal-oral and respiratory routes, although a few exceptions to the usual modes of transmission exist. Transmission of viral particles between humans occurs mainly via direct contact or fomites. Enterovirus RNA disappears over 2 to 5 weeks from the nasal cavity after an acute respiratory infection, and 97% of enterovirus-positive patients test negative for RNA by reverse transcription polymerase chain reaction (RT-PCR) 2 to 3 weeks after onset of illness (23). Enterovirus positivity, however, remained detectable longer than

culture positivity, suggesting that RNA may be remnants of past infections, without active replication.

Although the majority of enterovirus infections are subclinical, acute and clinical illnesses can present as including upper and lower respiratory illness, aseptic meningitis, encephalitis, acute haemorrhagic conjunctivitis, acute flaccid paralysis, myocarditis, and neonatal sepsis-like disease (24). Respiratory symptoms are mostly similar to a mild cold, yet in some cases can cause SARI requiring hospitalisation and may even be fatal (25). Enteroviruses are detected throughout the year, but have some peak activity in summer and autumn (26). Molecular techniques have demonstrated that new enterovirus genotypes causing respiratory illness have emerged or re-emerged (such as enterovirus-D68 [EV-D68], enterovirus 104, enterovirus 109, enterovirus 117 and coxsackie virus A21) (27–31). A study from Thailand in 2012 detected 11 different enterovirus genotypes from three species including EV-A (coxsackie virus A5, A6, A8, A10 and EV71), EV-B (coxsackie virus A9, B1, B2, echovirus 7 and echovirus 16) and EV-C (coxsackie virus A21) in paediatric patients with ILI (32).

No studies at the time of writing this thesis had determined whether different genotypes of enteroviruses are associated with mild (ILI) or severe acute respiratory illness (SARI) in other countries, and especially in South Africa, a country with high HIV prevalence.

1.3.2.2 Enterovirus D68

Recent reports of re-emerging lineages causing severe respiratory illness, some with fatal outcomes, were associated with EV-D68 infection in the Unites States of America (USA) (33). EV-D68 belongs to species EV-D and, until recently, sporadic outbreaks have been reported in Japan, Philippines, The Netherlands and the USA (34).

Reports of severe, fatal disease associated with EV-D68 infection in the USA renewed interest in enterovirus as the cause of SARI (33). EV-D68 infects primary target tissues of the respiratory tract directly, and occasionally infects the central nervous system (24). Bayesian phylogenetic analysis of EV-D68 strains globally, revealed the presence of three primary clades (A, B and C) although other literature have separated EV-D68 into clusters (1, 2 and 3) and lineages (1 and 2) (34–36). It is unclear if the recent upsurge in EV-D68 associated cases of pneumonia and flaccid paralysis-like illness are due to true changes in EV-D68 disease pathogenesis or improved molecular diagnostics (37). A previous study, using real-time reverse transcription polymerase chain reaction (qRT-PCR) assays, identified enteroviruses in 6% (515/8173) of patients that were hospitalised with pneumonia in South Africa (21). However, molecular characterisation of these viruses has not been done to determine whether some emerging lineages might be present in South Africa.

1.3.2.3 Echovirus 30

Enteroviruses are also a major cause of aseptic meningitis, and are usually associated with a favourable outcome. In temperate regions, seasonal aseptic meningitis usually peaks in summer and/or fall – similar to the seasonality observed in respiratory samples from South Africa (3,38). Many enterovirus genotypes circulate following an endemic pattern, but some genotypes such as echovirus 30 (E30) can cause widespread outbreaks. E30 is a member of enterovirus species B. The prototype strain of E30 (strain Bastianni) was isolated in 1958 (2).

E30 is one of the most common enterovirus genotypes responsible for sporadic as well as large-scale outbreaks of aseptic meningitis in many countries in Europe and North America (2,3). It is termed an epidemic virus since its circulation is highly variable from year to year (39). Phylogenetic studies with VP1 sequences indicate that E30 variants are continuously emerging and replacing the circulating

E30 strains, largely due to the error prone enterovirus RNA polymerase (31). Molecular epidemiology studies of E30 meningitis outbreaks show that point mutations give rise to substantial genetic diversity in the 1D gene encoding the VP1 capsid polypeptide (2). Twenty percent of codon sites in the 1D gene are subject to negative selection in most E30 population while a set of 10 codon positions were polymorphic and displayed signatures of relaxed negative selection (2). The substitution rate of E30 is 7.0 [95% high probability density, HPD 6.4-7.6] x 10^{-3} substitutions/site/year.

A common feature of E30 molecular epidemiology is progressive evolution of circulating lineages within one globally prevalent genotype (2,18). The dominant genotypes d-h (2) seem to be globally prevalent and all other lineages are seemingly extinct. E30 infection is usually cleared from the human host within a few days and, by analogy with another enterovirus infection - poliovirus, virus particles and virus infected cells are possibly cleared through the humoral response (15). The precise immune mechanisms involved in E30 infections have never been thoroughly investigated.

It has been reported that genotypes of E30 can become dominant for 3–4-year periods before they disappear from circulation (34), and they are therefore known as epidemic strains. E30 causes large outbreaks of aseptic meningitis that attract public health attention, and therefore there is a rich and extensive database of E30 VP1 sequences, from aseptic meningitis cases in the public domain. Despite the public health importance, relatively few complete genome sequences from sub-Saharan Africa are available for analysis (40). Only nineteen E30 full length genome sequences are currently available in GenBank, mainly isolated from aseptic meningitis cases with only two cases from respiratory illness cases from 2010, Thailand (https://www.ncbi.nlm.nih.gov/; accessed October 2018).

1.4 Rhinovirus species

1.4.1 Clinical manifestations and seasonality

Rhinoviruses are amongst the most commonly detected viruses in patients with upper respiratory tract infection and are considered to be associated, in most cases, with mild, self-limiting cold-like illnesses (41,42). The virus is usually mainly restricted to the upper respiratory airways and typically induces nasal congestion and rhinorrhoea, cough, sneezing, sore throat and malaise, with a spontaneous resolution within one to two weeks (43).

Rhinovirus infections are generally short-lived, and neutralizing antibodies that appear after viral clearance, have an essential role in protecting a person from reinfection. However, this humoral response seems to be genotype-specific with only little antibody cross-reaction (43). Rhinoviruses are detected throughout the year, with no evident seasonality (26). The large number of circulating genotypes supports the view that rhinoviruses do not circulate by waves or outbreaks of a given dominant genotype, which might explain the high frequency of reinfection during short periods (44).

1.4.2 Does rhinovirus contribute to the burden of severe respiratory illness?

Rhinoviruses have been identified from the nasopharynx of a large percent of healthy people without symptoms, and have not been considered as an important cause of acute respiratory hospitalisations in children (45). Recent advances in the development of sensitive molecular assays have considerably expanded the ability of laboratories to detect pathogens, especially viral agents. Together with the emergence of large pneumonia aetiology studies that include multiple viral causes, it has been shown that rhinovirus is not only associated with severe respiratory illness but also is identified frequently among asymptomatic patients (46,47).

Nonetheless, establishing a clinical association between pathogen detection and illness remains challenging since identification of a pathogen does not establish causality. A study performed in the USA on hospitalised children, estimated that

annual hospitalisation rates due to rhinovirus were 17.6/1000 in children less than 5 months followed by 6/1000 for children between 6-23 months and 2/1000 for children between the ages of 24-59 months. However when comparing children with or without a history of wheezing or asthma an increase in rhinovirus hospitalisations were observed 25.3/1000 vs. 3.1/1000, respectively (45).
In a previous study conducted in South Africa from 2012-2015, rhinoviruses were found to be significantly associated with mild and severe respiratory illness (21); whereas other studies could not establish any association (48–52). HIV infection is a known risk factor for increased severity of illness associated with the infection of several respiratory pathogens (7,53–55). In the South Africa study, 26% of enrolled patients with SARI were HIV infected and the HIV prevalence was 89% among patients aged 25-44 years (21). The significant association between rhinovirus detection and severe illness observed in this study but not in other studies raised the question whether the association was driven by the high HIV prevalence in South Africa.

1.5 Viral genomics, viral evolution and recombination events

Since enterovirus genotyping is based on partial VP1 sequencing, more extensive genome characterisation will assist in better understanding relationships between enterovirus genotypes since recombination frequently occurs in the P2 and P3 genomic regions (56). Complete genome sequencing is a powerful tool for pathogen characterization and are essential to describe circulating recombinant viruses (57).

In the last few years, complete genome sequences published in databases have grown exponentially along with the development of more sensitive and faster sequencing technology (58). Currently, next generation sequencing (NGS) is one of the most reliable and most popular technologies for this purpose. One of the strengths of NGS is ability to obtain detailed information about the pathogen in the sample in just one test in a relatively unbiased manner (57).

Viral evolution and recombination events can be studied using whole genome sequences. Enteroviruses evolve rapidly and even though most nucleotide substitutions are synonymous, amino acid changes do occur (59). Picornavirus RNA-dependent RNA polymerases are highly error prone and due to the absence of proof-reading activity, the substitution rate is high (1 mutation per 10^3 to 10^4 nucleotides) (60,61). Different parts of the enterovirus genome have distinct roles during the replication cycle and therefore may evolve differently and possibly exhibit remarkable independence during evolution (60). Mutation and recombination are the mechanistic alternatives for enterovirus evolution (60). The interaction between mutation (viral evolution) and recombination enables viruses to benefit from the advantages of having a few beneficial mutations without being affected by many deleterious mutations (40). Recombination occurs frequently in enteroviruses, thereby contributing to the genetic diversity and emergence of new genotypes. The first case of recombination in enteroviruses was originally reported in EV-B, specifically in polioviruses, and is since then recognised as a frequent and significant event occurring in enteroviruses (62). Recombination might occur within a specific genotype (intra-species) or between different genotypes of enteroviruses (inter-species)(40). To better understand enterovirus evolution, complete genome sequencing is important to identify inter and intra-typic recombination events, which might lead to emergence of strains with altered immune phenotypes, new tropisms, change in virulence or re-emergence of historic strains which may cause disease outbreaks and trigger serious public health concern (40,44,63).

The overall aim of this study was (i) to determine the epidemiology of enterovirus species by enterovirus type/ genotype and clinical syndrome during 2009-2014 and (ii) to determine the epidemiology of rhinovirus species and describe the fraction of rhinovirus infections attributable to illness, 2013-2015.

1.6 Aims and objectives

AIM 1: To determine the epidemiology and genotypes of enteroviruses associated with pneumonia in South Africa, 2009-2011

Objective 1: To determine the prevalence and seasonality of enterovirus in patients hospitalised with severe acute respiratory illness

Objective 2: To describe the clinical and epidemiological characteristics of patients hospitalised with severe acute respiratory illness testing positive for enterovirus

Objective 3: To determine the diversity of enterovirus genotypes present in patients hospitalised with severe acute respiratory illness by conducting traditional VP1 gene sequencing analysis

AIM 2: To compare enterovirus genotypes from patients with severe acute respiratory illness and influenza-like illness to those from asymptomatic controls in South Africa, 2012-2014

Objective 1: To determine the disease association, by HIV and age group, of enterovirus infection in patients with severe acute respiratory illness and influenza-like illness compared to asymptomatic controls

Objective 2: To determine the diversity of enterovirus genotypes present in patients with severe acute respiratory illness, influenza-like illness and asymptomatic controls by conducting traditional VP1 gene sequencing analysis

AIM 3: To perform genome-wide characterisation of most prevalent enterovirus, human echovirus 30, isolated from patients with severe acute respiratory illness in South Africa, 2013

Objective 1: To develop a viral enrichment sequence independent single primer amplification (SISPA) method to sequence the whole genome of enteroviruses

Objective 2: To conduct whole genome sequencing on echovirus 30 positive samples identified by traditional partial VP1 gene analysis in order to characterise viruses from respiratory illness cases across the complete genomes and to determine relatedness to strains reported elsewhere

AIM 4: To determine the epidemiology of rhinovirus species and to determine the attributable fraction of rhinovirus infection among HIV-infected and HIV-uninfected patients with mild and severe respiratory illness, South Africa, 2013-2015

Objective 1: To describe the clinical and epidemiological characteristics of patients hospitalised with severe acute respiratory illness and influenza-like illness testing positive for rhinovirus, compared to controls

Objective 2: To determine the attributable fraction of rhinovirus among patients hospitalised with severe acute respiratory illness and outpatients with milder influenza-like illness compared to asymptomatic controls.

Objective 3: To determine the adjusted prevalence corresponding to the prevalence of rhinovirus attributable to mild (influenza-like illness) and severe (severe acute respiratory illness), based on HIV status

Chapter 2

Materials and Methods

2.1 Description of the surveillance programmes and case definitions

2.1.1 Severe acute respiratory illness (SARI) surveillance

SARI surveillance is a prospective, hospital-based, surveillance programme conducted at two sites (Klerksdorp-Tshepong Hospital Complex, Klerksdorp, North West Province, and Edendale Hospital, Pietermaritzburg, KwaZulu-Natal Province) in South Africa (Figure 2.1).

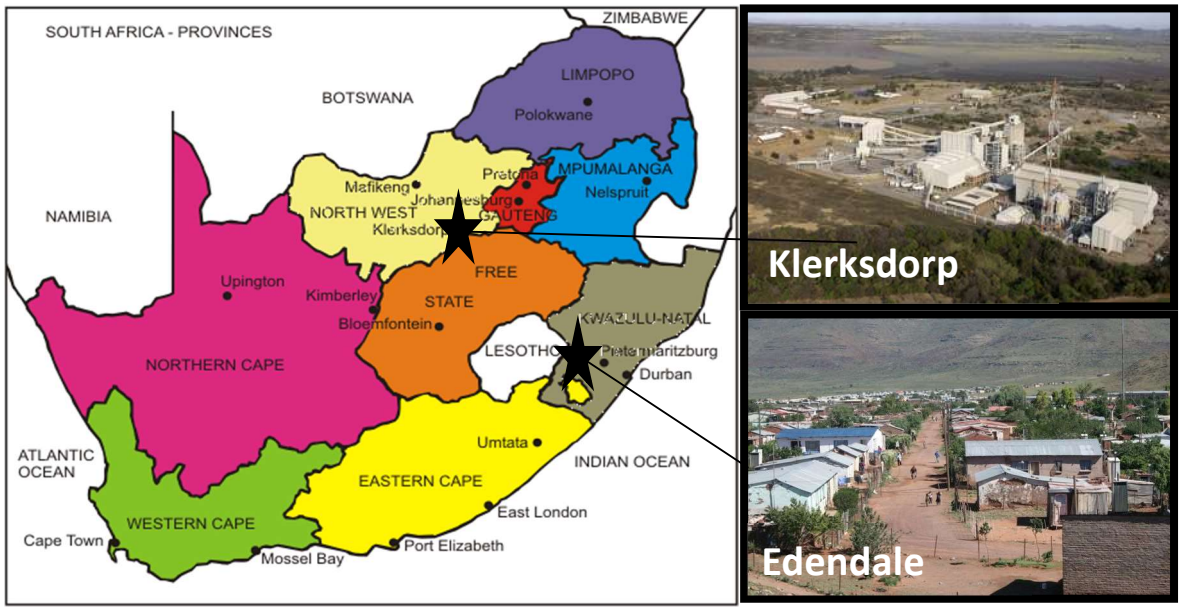


Figure 2.1: Map of provinces in South Africa – indicating the two surveillance sites (courtesy: Sibongile Walaza, Centre for Respiratory Diseases and Meningitis, NICD)

A case of SARI was defined as a patient who had symptom onset within 10 days prior to admission and who met age-specific clinical inclusion criteria as follows:

- A case in infants 2 days to less than 3 months of age included any patient with a diagnosis of suspected sepsis or physician-diagnosed acute LTRI, irrespective of signs and symptoms.
- A case in infants 3 months to less than 5 years of age included any patient with physician-diagnosed acute LRTI (e.g., bronchitis, bronchiolitis, and pneumonia) or pleural effusion.
- A case in individuals aged greater than and equal to 5 years of age included any patient with manifestation of acute LRTI with recorded temperature ≥38°C or history of fever and cough.

2.1.2 Severe chronic respiratory illness (SCRI) surveillance

A case of SCRI was defined as illness in a hospitalised person who had symptom onset greater than 10 days before admission and who met the age-specific clinical inclusion criteria described above (2.1.1).

2.1.3 Severe respiratory illness (SRI) surveillance

SRI cases include both SARI and SCRI cases. Therefore, a case of SRI was defined as a patient who was hospitalised (irrespective of symptom duration) and who met age-specific clinical inclusion criteria as follows:

- A case in infants 2 days to less than 3 months of age included any patient with a diagnosis of suspected sepsis or physician-diagnosed acute LTRI, irrespective of signs and symptoms.
- A case in infants 3 months to less than 5 years of age included any patient with physician-diagnosed acute LRTI (e.g., bronchitis, bronchiolitis, and pneumonia) or pleural effusion.

- A case in individuals aged greater than and equal to 5 years of age included any patient with manifestation of acute LRTI with recorded temperature ≥38°C or history of fever and cough.

2.1.4 Influenza-like illness (ILI) surveillance

ILI surveillance is a prospective, surveillance programme among patients presenting with ILI at two sites outpatient clinics (Jouberton Clinic, North West Province, and Edendale Gateway Clinic, KwaZulu-Natal Province) located in the same catchment area as the SARI and SCRI surveillance sites. An ILI case was defined as an outpatient of any age presenting with either a measured temperature greater than 38°C or, history of fever and cough with duration of symptom onset within 10 days.

2.1.5 Asymptomatic individuals

An asymptomatic individual or control was defined as persons of any age who sought medical care at one of the sentinel clinics for reasons other than fever, respiratory or gastrointestinal symptoms during the 14 days preceding the visit. Most controls visited the clinics for dental procedures, family planning, well-baby visits, voluntary HIV counselling and testing, or acute care for nonfebrile illnesses. The aim was to enrol one HIV-infected and one HIV-uninfected control every week in each clinic within each of the following age categories: 0–1, 2–4, 5–14, 15–54, and ≥55 years. Individuals were not followed-up to ensure they remained asymptomatic after enrolment.

2.2 Study procedures

Demographic and clinical characteristics were collected from all patients. Surveillance officers (registered/enrolled nurses) reviewed the clinical details of patients admitted to sentinel sites to establish whether cases met the surveillance

case definition/s. Identified cases were approached for inclusion in the sentinel surveillance. Consenting patients were asked to complete a questionnaire by interview on previous medical history and clinical data and to provide specimens.

Study staff completed case report forms for all enrolled controls, inpatients and outpatients. In addition, for inpatients, hospital records were reviewed to assess disease progression and outcome (i.e., discharge, transfer, or in-hospital death). Referral to hospital was recorded for all enrolled ILI cases. ILI cases that were referred to hospital were excluded from the analysis, as these patients did not adhere to the case definition of ILI anymore.

HIV results were obtained from a combination of two sources: (i) patient clinical records when available and (ii) for consenting patients, a dried blood spot was tested at the National Institute for Communicable Diseases (NICD) of the National Health Laboratory Service (NHLS), Johannesburg, South Africa. A HIV ELISA (Biorad Genscreen HIV ½ V2 EIA assay for screening and Biomerieux Vironostika HIV Ag/Ab EIA or DiaSorin Murex HIV Ag/Ab Combination EIA for confirmation) was used to test samples from patients aged ≥18 months, and PCR (COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Qualitative Test v2.0; Roche Molecular Systems, Inc., Branchburg, NJ) to test samples from children aged <18 months. Limited data were available on CD4+ T-cell counts on HIV positive patients.

2.3 Description of sample collection

Respiratory pathogen diagnosis depends on the collection of high-quality specimens, the rapid transport to the laboratory and appropriate storage before laboratory testing. Viruses are best detected in specimens containing infected cells and secretions.

Specimens should be collected as soon as possible, preferably during the first 72 hours (3 days) after onset of clinical symptoms in the acute phase of illness when

the greatest amount of viral shedding occurs, however specimens were processed even if collected greater than 3 days after the onset of symptoms.

For the optimal detection of respiratory pathogens, nasopharyngeal aspirates from children aged less than 5 years and combined nasopharyngeal-oropharyngeal swabs from persons aged greater than and equal to 5 years of age were collected from all enrolled participants. After collection, the specimens were placed in universal transport medium (Copan, Murruieta, California, USA), stored at 4–8°C, and transported within 72 hours of collection to the NICD for testing.

2.4 Laboratory methods and procedures

2.4.1 Total nucleic acid extraction

Total nucleic acid extraction for detection of enterovirus and rhinovirus was done using the MagNA Pure 96 instrument (Roche Diagnostics, Mannheim, Germany), which extracts 96 specimens per hour. Samples were prepared by adding 200 µl of each universal transport media in a processing cartridge. The extracted total nucleic acid (TNA) was eluted in 100 µl. The internal control was prepared by adding Poly (A) (final volume of 0.05 µg/ml) to the samples, in order to facilitate in the precipitation of RNA. The extraction was performed using the MagNA Pure 96 DNA and Viral Nucleic Acid Small Volume Kit together with the Viral Nucleic Acid Plasma Small Volume protocol.

Viral RNA for whole genome sequencing of enteroviruses was extracted from 200 µl viral cell culture supernatant using the total nucleic acid I kit and the Total Nucleic Acid Plasma_100_400 protocol on the MagNA Pure Compact instrument (Roche, Mannheim, Germany) according to manufacturer's instructions. The extracted total nucleic acid was eluted in 100 µl.

2.4.2 Laboratory diagnosis of enterovirus and rhinovirus

All respiratory specimens were tested for influenza A and B viruses, respiratory syncytial virus, parainfluenza virus types 1, 2 and 3, adenoviruses, human metapneumovirus, enteroviruses and rhinoviruses, using either a multiplex qRT-PCR assay (26) during 2013-2014 or the commercial Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) in 2015, with a Ct-value <37 regarded as a positive result.

For the multiplex qRT-PCR assay (26) on the Applied Biosystems 7500 Fast instrument (ThermoFisher, Waltham, Massachusetts, USA), complementary DNA (cDNA) was synthesized using the Transcriptor 1st Strand cDNA Kit (Roche Diagnostics), according to manufacturer's instructions. Each qRT-PCR reaction contained 2x LC Probe Master Mix (Roche, Mannheim, Germany), 1 µM of each primer, 0.5 µM of each probe, and 10 µl of cDNA reaction mixture as template, for a final volume of 30 µl. PCR cycles were initiated at 95°C for 15 minutes to activate thermostable DNA polymerase enzyme, followed by 45 cycles at 94°C for 15 seconds (denaturation), 60°C for 20 seconds (annealing), and 72°C for 10 seconds (elongation).

For the commercial Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) on the BioRad CFX96 instrument (Hercules, California, USA), each qRT-PCR reaction contained 5x Real-time One-step Buffer, 5x PR2 or PR3 primers/probes, 1U Real-time One-step enzyme and 8 µl of RNA as template, for a final volume of 25 µl. The initial step was reverse transcription at 50°C for 20 minutes followed by activation at 95°C for 15 minutes to activate thermostable DNA polymerase enzyme, followed by 45 cycles at 95°C for 10 seconds (denaturation), 60°C for 1 minute (annealing), and 72°C for 10 seconds (elongation).

The qRT-PCR used for enterovirus testing, is a sensitive method for the routine detection of all members of the enterovirus genus. This method has a weak cross-

reactivity to high-titre rhinovirus stocks, but not with rhinovirus-positive clinical samples (27).

2.4.3 Sanger sequencing of enterovirus VP1 gene region

A 375 base pair region targeting the 5' proximal part of VP1 region of enteroviruses were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of EV-68 (GenBank accession number AY426531). cDNA synthesis using random primers as well as oligo(dT) primers was performed with Transcriptor 1st Strand cDNA Kit (Roche Diagnostics, Mannheim Germany).

First round and nested PCR was performed as per Nix *et al*., 2006 (64). 5 µl of the cDNA reaction mixture was used in the first round PCR (final volume, 25 µl), consisting of 2x AmpliTaq Gold 360 PCR buffer (ThermoFisher, Waltham, Massachusetts, United States) and 10 pmol each of primers 224 (5'-GCI ATG YTI GGI ACI CAY RT-3') and 222 (5'-CIC CIG GIG GIA YRW ACA T-3'). PCR cycles were initiated at 95°C for 10 minutes to activate thermostable DNA polymerase enzyme, followed by 40 cycles at 95°C for 30 seconds (denaturation), 42°C for 30 seconds (annealing), and 72°C for 60 seconds (elongation). A final elongation step was performed at 72°C for 7 minutes.

2 µl of the first round PCR was added to a nested PCR, consisting of 2x AmpliTaq Gold 360 PCR buffer (ThermoFisher, Waltham, Massachusetts, United States) and 10 pmol each of primer AN89 (5'-CCA GCA CTG ACA GCA GYN GAR AYN GG-3') and AN88 (5'-TAC TGG ACC ACC TGG NGG NAY RWA CAT-3'), in a final volume of 25 µl. PCR cycles were initiated at 95°C for 10 minutes to activate Taq DNA polymerase enzyme, followed by 40 cycles at 95°C for 30 seconds (denaturation), 60°C for 30 seconds (annealing), and 72°C for 60 seconds (elongation). A final elongation step was performed at 72°C for 7 minutes.

5 µl of the nested PCR product (375 bp) were separated and visualized on a 1% agarose gel, stained with 2x GelStain Red (Biotium, Fremont, California, USA) and run for 1 hour at 120V using a 100 bp molecular weight marker (Roche, Mannheim, Germany) as a size reference. Slight variations in the sizes of the PCR products (350 to 400 bp) were observed due to VP1 gene length differences in the different genotypes.

Amplicons were purified using the ExoSAP-IT enzyme system (USB Corporation, Cleveland OH, USA). The exonuclease I and the Shrimp Alkaline Phosphatase were mixed in a 1:1 ratio and 3 µl of the mixture were added to 10 µl of amplicons. The mix were incubated at 37°C for 15 minutes followed by 80°C for 15 minutes.

Purified amplicons were sequenced using the Big Dye terminator version 3.1 cycle Sequencing Ready Reaction kit (Life Technologies, Foster City, USA) by using 10 pmol primers AN89 and AN88. 2 µl of the purified amplicon was used (final volume, 10 µl), consisting of 5x BigDye buffer, 5x BigDye mix and 10 pmol each of primer in the forward and reverse reaction, respectively. PCR cycles were initiated at 96°C for 0 seconds, followed by 25 cycles at 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 seconds.

After the sequencing reaction, the amplicons were cleaned using the BigDye Xterminator Purification kit (Life Technologies, Foster City, USA). 2 µl of 2.2 % SDS were added to 10 µl of cycle sequenced amplicons and incubated at 98°C for 5 minutes, 25°C for 10 minutes, and 15°C for 7 minutes, to reduce secondary structures. 10 µl X-terminator mix and 55 µl SAM solution were added to the SDS-treated amplicon. The reactions were shaked at 1800 rpm for 30 minutes and centrifuged at 1000 g for 2 minutes.

2.4.3.1 Assembly and phylogenetic analysis of enteroviruses partial VP1

Sequences were assembled using Sequencher® version 5 (Gene Codes Corporation, Michigan, USA). Sequence alignments were performed using MAFFT multiple sequence alignment program (65). The Kimura-2 parameter nucleotide substitution model determined as the optimal substitution model by using jModelTest was used for the Neighbor-joining (NJ) analysis (66,67) and the NJ trees were generated using MEGA 5.2 software (68,69).

2.4.4 Enterovirus isolation

Human epithelial type 2 (HEp-2) cells were used for expansion and passage of the E30 positive samples. Cells were cultured in DMEM (Dulbecco's Modified Eagle's Medium) containing 10% fetal bovine serum in an incubator at 37°C with 5% CO_2. Most cultures underwent two to four passages before being subjected for whole genome sequencing. Positive cultures were identified through characteristic cytopathic effect as seen under a light microscope. The virus was harvested by scraping infected cell layers and mixed with culture fluid by vortexing, and the suspension was subsequently centrifuged at 600 g for 10 minutes at 4°C to remove cellular debris.

2.4.5 Enterovirus genome sequencing

2.4.5.1 Viral enrichment and sequence independent single primer amplification (RACE-SISPA) of echovirus 30

2.4.5.1.1 Viral enrichment

The E30 culture supernatant fluids were filtered with a 0.22 micron filter, to remove bacteria and larger cellular components. To eliminate residual cellular DNA contaminants, the filtrate was treated with 0.1U/µl DNase I and incubated at

37°C for 1 hour. Viral RNA was extracted from 200 µl virus-enriched cell culture supernatant using the total nucleic acid isolation I kit on the MagNA Pure Compact instrument (Roche, Mannheim, Germany) according to manufacturer's instructions, and the RNA was eluted in 100 µl volumes. The eluates were treated again with 0.1U/µl DNase I.

A Nanodrop spectrophotometer (ThermoFisher, Waltham, Massachusetts, United States) was used to determine DNA concentrations and concentration >0.5 ng/µl was deemed sufficient for NGS. A bioanalyzer (Agilent technologies, California, United States) was used to assess the quality of the starting material and monitoring the size distribution after fragmentation and adapter ligation.

2.4.5.1.2 Rapid amplification of cDNA (RACE)

Poly-U RNA was generated by addition of 7 µl of the extracted RNA to a reaction mixture containing 10x NEBuffer 2 (New England Biolabs, Ipswich, Massachusetts, USA), 10 µM UTP, 2 units/µl poly-U polymerase and 40 units/µl RNase inhibitor (Promega, Madison, Wisconsin, USA) in a total volume of 20 µl. The reaction tubes were incubated at 37°C for 10 minutes and cooled to 4°C.

2.4.5.1.3 Sequence independent single primer amplification (SISPA)

SISPA was performed as per method describe by Djikeng *et al.,* 2008 (70) with slight modifications. First-strand cDNA synthesis using SISPA primer (10 µM), FR26RV-N (5'-GCC GGA GCT CTG CAG ATA TCN NNN NN-3'), as well as an anchored primer (1 µM), FR40RV_anchored A (5'-GCC GGA GCT CTG CAG ATA TCA AAA AAA AAA AAA VN-3'), was performed using Thermo Scientific Maxima H minus Reverse Transcriptase (ThermoFisher, Waltham, Massachusetts, United States). The Poly-U RNA (7.5 µl) was first denatured together with the SISPA (0.4 µM) and anchored primers (0.2 µM) as well as 2%

DMSO in a total volume of 12 µl for 5 minutes at 65°C, cooled to 4°C for 2 minutes and then moved to ice immediately.

The first-round cDNA synthesis reaction was completed by addition of the following reagents to the 12 µl denatured RNA/primer mix: 5 x RT Buffer, 200 units/µl Maxima H minus reverse transcriptase, 5 µM Venter SISPA primer (5'-GCC GGA GCT CTG CAG ATA TCG GCC ATT ATG GCC (Ribo-GGG)3') (70), dNTP mix (0.5 mM) (Promega, Madison, Wisconsin, USA), 40 units/µl RNase inhibitor (Promega, Madison, Wisconsin, USA) in a total volume of 20 µl. The reaction was subjected to the following conditions: 10 minutes at 25°C, 30 minutes at 50°C, 15 minutes at 65°C, 5 minutes at 85°C and cooled to 4°C.

Second-strand cDNA synthesis was done on each first strand cDNA product by adding 1 µl RNAse H (Ambion) and 1 µl Klenow large fragment DNA polymerase with 3'-5' exonuclease activity; and the reaction mix subjected to the following conditions: 60 minutes at 37°C, 15 minutes 75°C and cooled to 4°C. The double stranded cDNA products were purified using Qiagen MinElute PCR purification kit (Hilden, Germany) as per manufacturer's instructions. All samples were eluted with 12 µl elution buffer.

Random fragment amplification was done in a 50 µl reaction using 10 µl of double stranded cDNA as template together with 2x MyTaq Red mix, 10 µM primer FR20RV (5'-GCC GGA GCT CTG CAG ATA TC-3') (70) and nuclease-free water. The reaction was subjected to 98°C for 30 seconds, 40 cycles of 98°C for 10 seconds, 54°C for 20 seconds, 72°C for 45 seconds followed by a final extension of 72°C for 5 minutes and cooled to 4°C. A second shorter amplification was done. To each 50 µl sample reaction, the following was added: 12.5 µl 2x MyTaq Red mix, 1 µl primer FR20RV (10 µM) and 13.5 µl nuclease-free water. The whole mixture was subjected to 98°C for 30 seconds, 15 cycles of 98°C for 10 seconds, 54°C for 20 seconds, 72°C for 45 seconds followed by a final extension of 72°C for 5 minutes and cooled to 4°C.

The DNA products were purified using Qiagen MinElute PCR purification kit (Hilden, Germany) as per manufacturer's instructions. All samples were eluted with 12 µl elution buffer. The concentration of the DNA was quantified using Qubit® Fluorometer (Invitrogen, Carlsbad, California, United States) using manufacturer's instructions.

2.4.5.2 Next generation sequencing on the MiSeq platform

Viral genomes were sequenced with the Illumina MiSeq platform (Illumina, California, USA). Libraries were constructed using 1 ng of DNA with the Nextera XT DNA library Preparation kit (Illumina, California, USA). All clean-up steps were performed using Agencourt AMPure XP beads (Beckman Coulter, California, USA). The libraries were quantified and quality controlled using the Agilent High Sensitivity DNA Kit (Agilent technologies, California, USA). Cluster generation and sequencing were completed utilizing Illumina's standard protocol (71).

2.4.5.3 Data analysis: Trimming, mapping of the reads and assembly

Reads were de-multiplexed following the removal of tags and adaptors. Following importation in CLC Genomics Workbench (Qiagen, Hilden, Germany), reads were paired, trimmed and merged using the following parameters: Trim quality score limit = 0.01; Trim ambiguous nucleotides with maximum number of ambiguities = 1. After trimming, reads shorter than 50 nucleotides were discarded. For E30 the sequences were then mapped against the reference sequence (Human E30 strain Bastianni – GenBank accession number AF311938) using the following parameters: Mismatch cost = 10; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.5; Similarity fraction = 0.95. The consensus sequence was extracted and exported for further analysis.

2.4.6 Phylogenetic analysis of echovirus 30 whole genomes

Phylogenetic trees were constructed by the neighbour-joining method implemented in the MEGA 7.0.26 program (72) using the Kimura 2-parameter model. Regions containing alignment gaps were ignored in the analysis. The branch lengths of the dendrogram were determined from the topologies of the trees and were obtained by majority rule consensus among 1000 bootstrap replicates. Bootstrap values greater than 75% were considered statistically significant for grouping. Each gene segment was analysed with Basic Local Alignment Search Tool (BLAST) to confirm that no inter-typic recombination occurred (73). Similarity plot and boot scanning analyses were performed using the RDP4 software and DataMonkey (74–76). For similarity plot analyses, a 200-nucleotide window was moved in 20-nucleotide steps, and boot-scanning analyses were run with the neighbor-joining method. GARD (Genetic Algorithm for Recombination Detection) is a method to screen a multiple sequence analysis for the presence of recombination and is extremely useful as a pre-processing step for selection inference. Because recombinant sequences cannot be adequately described with a single phylogenetic history, selection inference on recombinant data often leads to a significant increase in false positives. GARD alleviates this concern by comprehensively screening an alignment for recombination breakpoints and inferring a unique phylogenetic history for each detected recombination block (76).

2.5 Statistical data analysis

Unconditional exact logistic regression was used to assess the association of enterovirus species with disease severity among patients with mild (ILI) or severe illness (SARI) using asymptomatic individuals as control group. Exact logistic regression was used to account for the fact that no EV-D species was detected among controls. EV-B was used as the comparison group as it was the species most frequently detected.

In addition, unconditional penalized logistic regression were used to estimate the attributable fraction (AF) of rhinovirus-associated hospitalisation and outpatient consultation by comparing the rhinovirus detection rate (observed) among ILI, SARI and SCRI cases to those of controls. Penalized logistic regression was selected to account for separation problems (i.e. no observations in one of the comparison cells) (77). The analysis was implemented among any inpatient with SRI (i.e., patients with SARI or SCRI combined). The AF was estimated from the odds ratio (OR) obtained from the logistic regression models as follows:

$$AF = \frac{OR - 1}{OR} * 100$$

Subsequently, the rhinovirus percentage positive associated with illness among patients with ILI, SARI, SCRI and SRI ($Prev_{Illness}$) from the observed percentage positive ($Prev_{Observed}$) was estimated as follows: $Prev_{Illness} = Prev_{Observed} * AF$
The analysis was implemented overall, and stratified by HIV serostatus within the following age categories: <1 year, 1-4 years, 5-24 years, 25-44 years, 45-64 years, ≥65 years as well as <5 years and ≥5 years. These age groups were selected for consistency and comparability with previous studies from South Africa.

All estimates were adjusted for co-infections with the other respiratory viruses investigated in this study and any underlying medical condition (treated as dichotomous variable indicating presence or absence of any evaluated condition) as they were significantly associated with the outcome. This was obtained by including covariates for any underlying medical condition and positivity for the individual viruses investigated in the study in the multivariable penalized logistic regression model.

Underlying medical conditions were asplenia, including asplenia or sickle cell anemia; chronic illness, including chronic lung, renal, liver, or cardiac disease, diabetes mellitus and asthma; other immunocompromising conditions (excluding HIV), including organ transplant, primary immunodeficiency, immunotherapy, and malignancy; neurologic disorders; burns; obesity; malnutrition; and

prematurity. Prematurity and malnutrition were included as underlying medical conditions only in the analyses implemented among children aged <1, 1-4 and <5 years. HIV infection and age (i.e., <1, 1–4, 5–24, 25–44, 45–64, ≥65) were also included as covariates in the non HIV-stratified overall model and the models implemented among children aged <5 (<1 and 1-4 years) and person aged ≥5 years (5-24, 25-44, 45-64 ≥65 years).

The Chi-squared or Fisher's exact test were used for comparison of categorical variables. Significant associations were considered at $p<0.05$. The statistical analysis was -implemented using STATA version 14.1 (StataCorp, College Station, Texas, USA).

Chapter 3

Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients with Severe Acute Respiratory Illness, 2009-2011

Reference:

Hellferscee O, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, Cohen AL, Pretorius M, Cohen C, Madhi SA, Venter M. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza and Other Respiratory Viruses. 2017 May; 11(3):211-219

3.1 Introduction

Several respiratory viruses, including enterovirus have been associated with SARI in South Africa (21). Although the majority of enterovirus infections are subclinical; they can cause a variety of acute clinical syndromes including upper and lower respiratory tract illness (24).

A previous study identified enteroviruses in 6% (515/8173) of patients that were hospitalised with pneumonia in South Africa (26). However molecular characterisation of these viruses has not been done to determine whether some emerging clades might be present in South Africa. The contribution of specific clades to severe disease in light of the high HIV sero-prevalence in the country, were not investigated.

The aim was to describe the epidemiology of enteroviruses and determine which genotypes were circulating among South African patients hospitalised with SARI during 2009-2011.

3.2 Materials and Methods

3.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI surveillance programme (described in section 2.1).

3.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients as described in section 2.2 and 2.3.

Specimens were tested for the presence of 10 respiratory viruses (enteroviruses, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (26) (section 2.4.1-2.4.2).

The study period included samples collected from January 2009 to December 2011 (n=10,260). Due to challenges with availability of reagents, we did not test for adenovirus from August–October 2011. For this period, retrospective testing was done on 30% of randomly selected samples and the prevalence extrapolated. Enterovirus positive specimens from each year (2009: n=75/315, 2010: n=48/200 and 2011: n=77/327) were randomly selected for genotyping and molecular characterization to obtain 24% of all enterovirus positive samples per study year.

3.2.3 Amplification and sequencing of enteroviruses partial VP1

The 5' proximal part of VP1 region were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (see methods section 2.4.3).

3.2.4 Statistical analysis

Differences in characteristics of enterovirus genotypes were assessed using the Fisher's exact test. P-values <0.05 were considered to be statistically significant. Analysis was performed using STATA 13 (Stata Corporation, Texas USA).

3.3 Results

3.3.1 Clinical and epidemiological characteristics of patients hospitalised with enterovirus-associated SARI

During the study period, enterovirus was detected in 8.2% (842/10,260) of SARI cases. For observations with complete data, as displayed in Table 3.1, the majority of enterovirus -positive patients (86.5%, 726/839) were <5 years of age. Enteroviruses were detected throughout the study, although the detection rate was lower in 2010 (5.8%) than in 2009 (10.1%) and 2011 (8.9%, p=0.047). Sixty-three percent (501/795) of enterovirus cases had co-infections with other respiratory viruses and bacteria. Respiratory syncytial virus and adenovirus (21.5% and 30.6%, respectively) were the most frequently detected viral co-infections with enterovirus. The majority of patients (70.0%, 582/831) had a recorded fever and 25% (211/829) of cases had symptoms of ≥3 days prior to admission. Almost all enterovirus associated SARI cases (93.4%, 781/836) had antibiotics prescribed on admission and the majority (74.5%, 618/829) were hospitalised for less than three days.

The in-hospital case fatality rate was 2.2% (19/840). However, these deaths could not be attributed solely to enterovirus infection since these patients were also co-infected with various other pathogens: 15.8% (3/19) *Mycobacterium tuberculosis*;

11.8% (2/17) adenovirus; 10.5% (2/19) human metapneumovirus; 15.8% (3/19) parainfluenza virus 2; 10.5% (2/19) parainfluenza virus 3; 15.8% (3/19) respiratory syncytial virus and 5.3% (1/19) *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Streptococcus pneumoniae* (in blood samples).

HIV status was known for 64% (541/842) of cases, among whom the prevalence of HIV was 19.8% (107/541). Close to 62% (8/13) of the cases with fatal outcome was HIV positive. The HIV status among all SARI patients were known for 79% (8120/10260) of cases, among whom the prevalence of HIV was 49.52% (4021/8120), and the case fatality rate 5.6% (577/10260).

3.3.2 Molecular epidemiology of enterovirus species A to D

In this study 48 (24%) from the 200 randomly selected enterovirus positive viruses could be genotyped of which 3 grouped with the rhinoviruses; these were excluded from further analysis but are indicated as reference strains on the phylogenetic tree (Figure 3.1). EV-B (44%, 20/45) and EV-D (22%, 10/45) species were more prevalent than EV-A (16%, 7/45) and EV-C (18%, 8/45) with bootstrap support of 95-99% for clusters (Figure 3.1).

Enteroviruses were detected throughout 2009-2011, with no evident seasonality, although a higher EV-D68 (only EV-D species detected) activity was seen in 2011 (Figure 3.2A) whereas enterovirus species B, C and D co-circulated during 2009-2011 (Figure 3.2B) and EV-A was detected only in 2009 and 2010.

The majority of EV-A (86%, 6/7) and EV-B (55.0%, 11/20) strains were detected in 2010, while EV-C (62.5%, 5/8) and EV-D (70.0%, 7/10) were more commonly detected in 2011 (Table 3.1). The majority of patients infected with EV-B (80%, 16/20) and EV-D (80%, 8/10) had a recorded fever. A high proportion of viral co-infections (other than enterovirus) were found among EV-B (62.5%, 10/16) and EV-D (50%, 5/10) positive patients. All HIV-infected patients in this subset tested, were infected with EV-B-strains (Table 3.1, Figure 3.1).

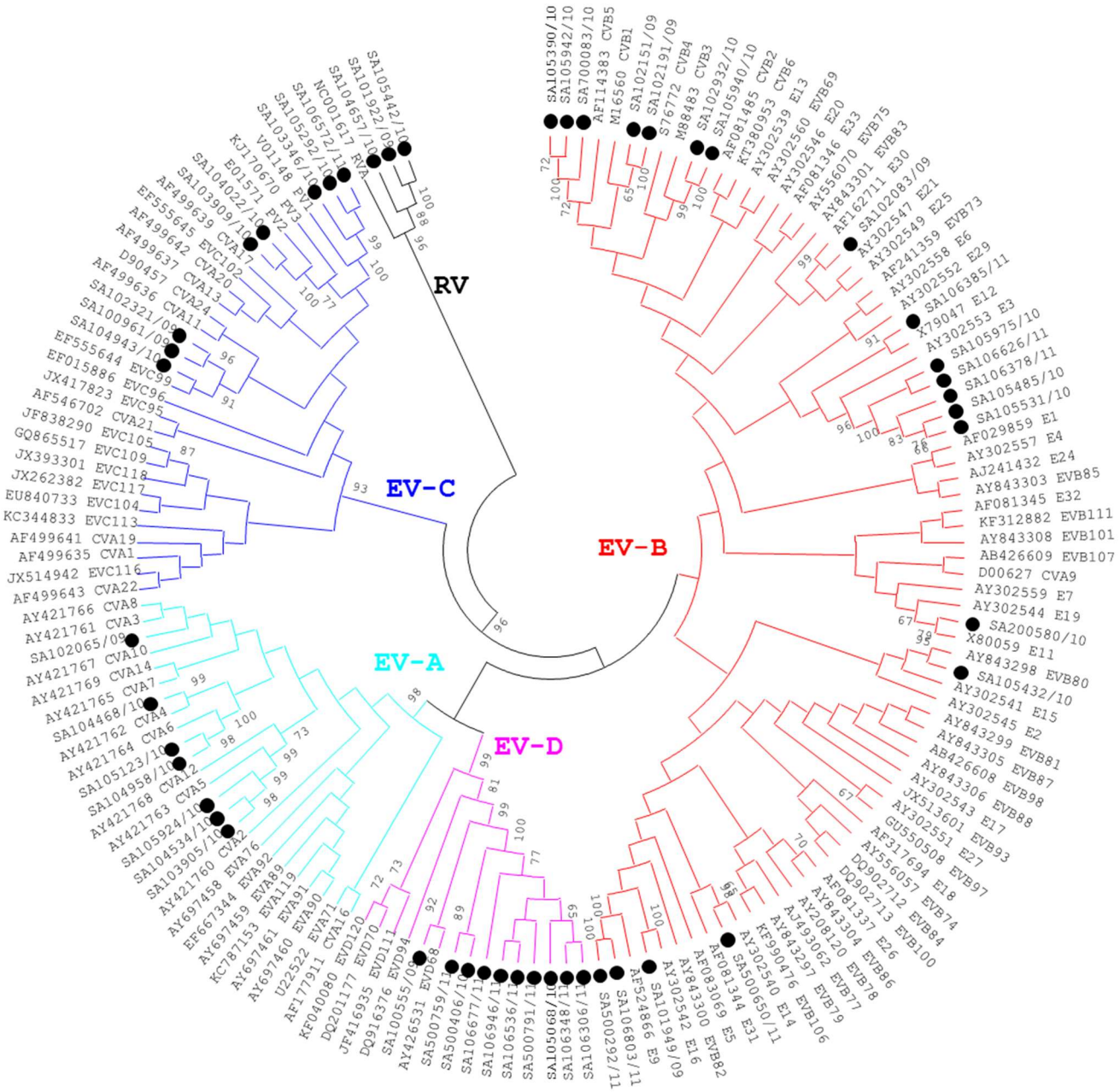
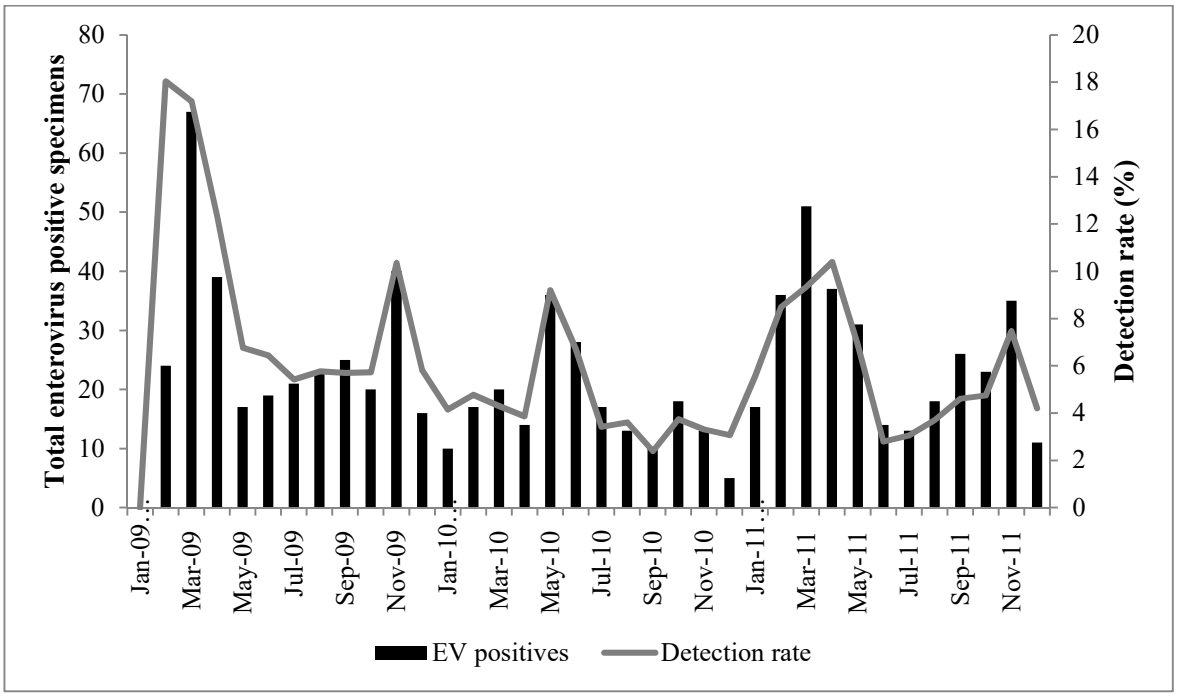


Figure 3.1: Molecular characterization of enterovirus clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implemented in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Black dots = South African samples 2009-2011.

A



B

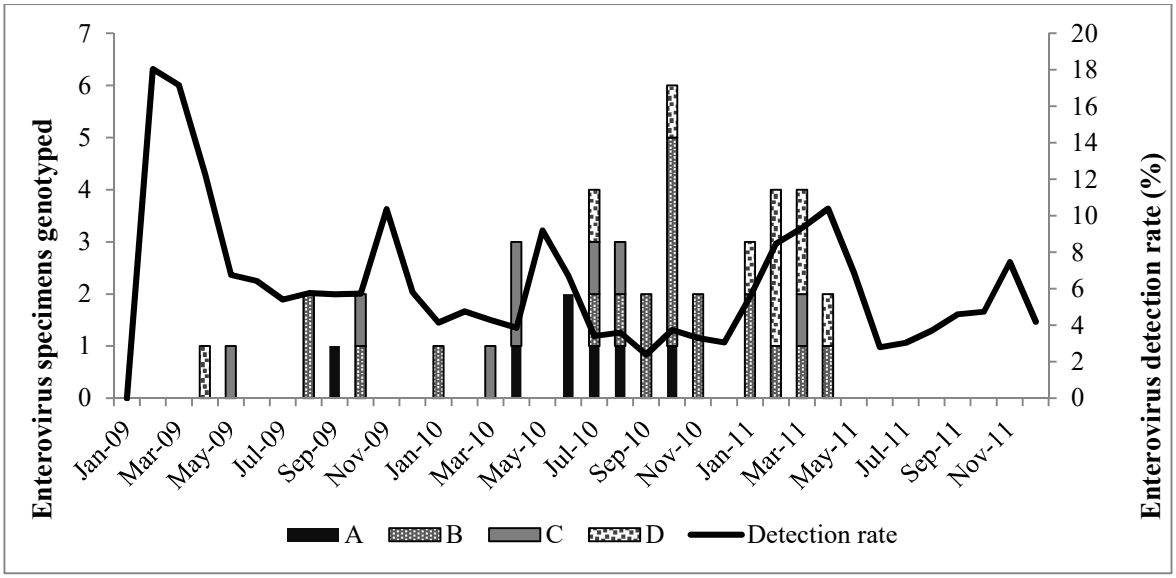


Figure 3.2: Epidemiologic graphs showing: (A) Distribution of enterovirus in South Africa during 2009-2011 and (B) Distribution of enterovirus species detected in South Africa during 2009-2011

3.3.3 Enterovirus genotypes and circulation of EV-D68 clades during 2009-2011

Seventeen distinct enterovirus genotypes were identified of which EV-D68 (22%, 10/45) and echovirus 3 (E3) (11%, 5/45) were the most frequently detected. Other genotypes detected include coxsackie virus A (3, 4, 5, 6), coxsackie virus B (1, 3, 5), echovirus (9, 11, 12, 14, 16, 30, 80, 99) and poliovirus (1, 2) (Figure 3.1).

Here EV-D68 strains are grouped according to phylogenetic clades A, B and C. The majority of EV-D68 strains (90%, 9/10) are in clade B. Previously reported South African EV-D68 strains from 2000-2001 (7) clusters in clades C and A (Figure 3.3). The South African clade B strains clustered with strains from respiratory specimens from Japan, 2010 and the Netherlands, 2009-2010 (8/10) and from the Philippines, 2008 and 2011 (1/10). The South African clade A strain clustered with strains from respiratory specimens from Italy, 2008, the Netherlands, 2008, the Philippines, 2011 and Kenya, 2010.

The designated clades B1 and B3, described recently were based on more complete gene sequence data (~930 base pairs for complete VP) (9, 10) and therefore the phylogenetic tree (based on 300 base pair VP1 fragment) did not have good bootstrap support for clade B1 and B3. Demographic and clinical characteristics of patients in whom EV-D68 was isolated are described in Table 3.2. Ages ranged from 3 to 54 months and respiratory symptoms included cough, chest indrawing, breathing difficulty and tachycardia.

No amino acid changes in the VP1 region that could lead to increased virulence (21) were observed in the South African samples (results not shown).

Table 3.1: Clinical and epidemiological characteristics of patients hospitalised with enterovirus associated severe acute respiratory illness at surveillance sites, South Africa, 2009-2011

Variables	All SARI samples n/N (%)	All enterovirus positive samples n/N (%)	Enterovirus genotyped n/N (%)	Species A n/N (%)	Species B n/N (%)	Species C n/N (%)	Species D n/N (%)	p value
Demographic and clinical characteristics								
Age (in years)								0.836
<1	4263/13598 (31.4)	377/839 (44.9)	23/45 (51.1)	3/7 (42.9)	10/20 (50)	5/8 (62.5)	5/10 (50)	
1-4	2411/13598 (17.7)	349/839 (41.6)	19/45 (42.2)	4/7 (57.1)	8/20 (40)	2/8 (25)	5/10 (50)	
≥5	6924/13598 (50.9)	113/839 (13.5)	3/45 (6.7)	0/7 (0)	2/20 (10)	1/8 (12.5)	0/10 (0)	
Male Sex	6640/13602 (48.8)	489/839 (58.2)	21/45 (46.7)	4/7 (57.1)	9/20 (42.9)	3/8 (37.5)	5/10 (50)	
Year								0.047
2009	3679/13656 (26.9)	315/3128 (10.1)	8/45 (17.8)	1/7 (14.3)	4/20 (20)	2/8 (25)	1/10 (10)	
2010	4613/13656 (33.8)	200/3460 (5.8)	24/45 (53.3)	6/7 (85.7)	11/20 (55)	5/8 (62.5)	2/10 (20)	
2011	5364/13656 (39.3)	327/3672 (8.9)	13/45 (28.9)	0/7 (0)	5/20 (25)	1/8 (12.5)	7/10 (70)	
Symptoms ≥3 days prior to admission	6867/13421 (51.2)	211/829 (25.5)	9/45 (20)	1/7 (15.3)	7/20 (35)	1/8 (12.5)	0/10 (0)	0.125
Fever (≥38˚C)	10755/13547 (78.4)	582/831 (70)	31/45 (68.9)	3/7 (42.9)	16/20 (80)	4/8 (50)	8/10 (80)	0.158

Oxygen therapy	4491/13508 (33.3)	236/831 (28.4)	15/45 (33.3)	2/7 (28.6)	6/20 (30)	1/7 (12.5)	6/10 (60)	0.215
Antibiotics prescribed on admission	12903/13549 (95.2)	781/836 (93.4)	41/45 (91.1)	7/7 (100)	17/20 (85)	8/8 (100)	9/10 (90)	0.809
Hospitalisation (≥ 3 days)	9053/13482 (67.2)	362/832 (43.5)	18/45 (40.9)	1/7 (14.3)	11/20 (55)	2/8 (25)	4/10 (44.4)	0.236
In-hospital death	755/13526 (5.6)	19/840 (2.2)	2/45 (4.4)	0/7 (0)	1/20 (5)	1/8 (12.5)	0/10 (0)	0.606
Underlying medical conditions								
HIV	5264/10666 (49.4)	107/541 (19.8)	4/34 (11.8)	0/5 (0)	4/17 (23.5)	0/3 (0)	0/9 (0)	0.412
Prematurity[a]	178/6661 (2.6)	21/725 (2.9)	2/45 (4)	0/7 (0)	1/20 (5)	1/8 (12.5)	0/10 (0)	0.606
Other medical conditions[b]	1189/13588 (8.8)	62/838 (7.4)	4/45 (8.9)	0/7 (0)	3/20 (15)	1/8 (12.5)	0/10 (0)	0.586
Co infections								
Streptococcus pneumoniae (blood)	706/10242 (6.9)	23/505 (4.6)	2/31 (6.5)	0/7 (0)	2/14 (14.3)	0/4 (0)	0/9 (0)	0.729
Respiratory syncytial virus	1897/13413 (14.1)	181/842 (21.5)	5/45 (11.1)	0/7 (0)	2/20 (10)	1/8 (12.5)	2/10 (20)	0.773
Para-influenza 2 virus	148/13413 (1.1)	20/842 (2.4)	2/45 (4.4)	0/7 (0)	2/20 (10)	0/8 (0)	0/10 (0)	0.798
Influenza virus (A and B)	1168/13430 (8.7)	33/842 (3.9)	2/45 (4.4)	0/7 (0)	2/20 (10)	0/8 (0)	0/10 (0)	0.798
Adenovirus	1129/12228 (15.8)	236/772 (30.6)	9/39 (23.08)	1/6 (16.7)	6/16 (37.5)	0/7 (0)	2/10 (20)	0.246
Any viral coinfection	7052/12860 (54.8)	501/795 (63)	19/39 (48.7)	3/6 (50)	10/16 (62.5)	1/7 (14.3)	5/10 (50)	0.205

Abbreviations: ICU: Intensive Care Unit; HIV: human immunodeficiency virus.

[a] Prematurity was classified as birth before 37 weeks of gestation as reported on the road-to-health card.

[b] Other evaluated underlying medical conditions included: asplenia or sickle cell anaemia; chronic illness, including chronic lung, renal, liver, or cardiac disease and diabetes; other immunocompromising conditions (excluding HIV), including organ transplant, primary immunodeficiency, immunotherapy and malignancy; neurological disorders and burns.

Table 3.2: Demographic and clinical characteristics of ten patients in whom EV-D68 was identified in South Africa

Characteristics	SA500759/11	SA106306/11	SA105068/10	SA106946/11	SA106677/11	SA106348/11	SA500406/10	SA500791/11	SA106536/11	SA100555/09
Demographic characteristics										
Age (in months)	25	15	3	11	3	16	54	16	11	5
Sex	F	F	F	M	M	F	M	M	F	M
Year	2011	2011	2010	2011	2011	2011	2010	2011	2011	2009
Medical history										
Underlying illness	N	N	N	N	N	N	N	N	N	N
HIV infected	N	N	N	N	N	N	N	N	UNK	N
Clinical presentation and course										
Symptoms (days) prior to admission	1	1	1	1	0	1	2	1	1	0
Fever (≥38˚C)	Y	Y	Y	Y	N	Y	Y	Y	Y	UNK
Respiratory symptoms	C, B, S, L, A	C, B, CI, T	C, B, T, A, V	C, T	C, CI	C, B, T	C, B, CI, S	C, B, T	C, B, CI, S, T, A, V	C, B, CI, T

Abbreviations: M: Male; F: Female; N: No; UNK: Unknown; Y: Yes; C: Cough; S: Stridor; R: Rhinorrhoea; M: Malaise; F: Fatigue; A: Anorexia; L: Lethargy; B: Breathing difficulty; CI: Chest indrawn; T: Tachycardia; V: Vomiting

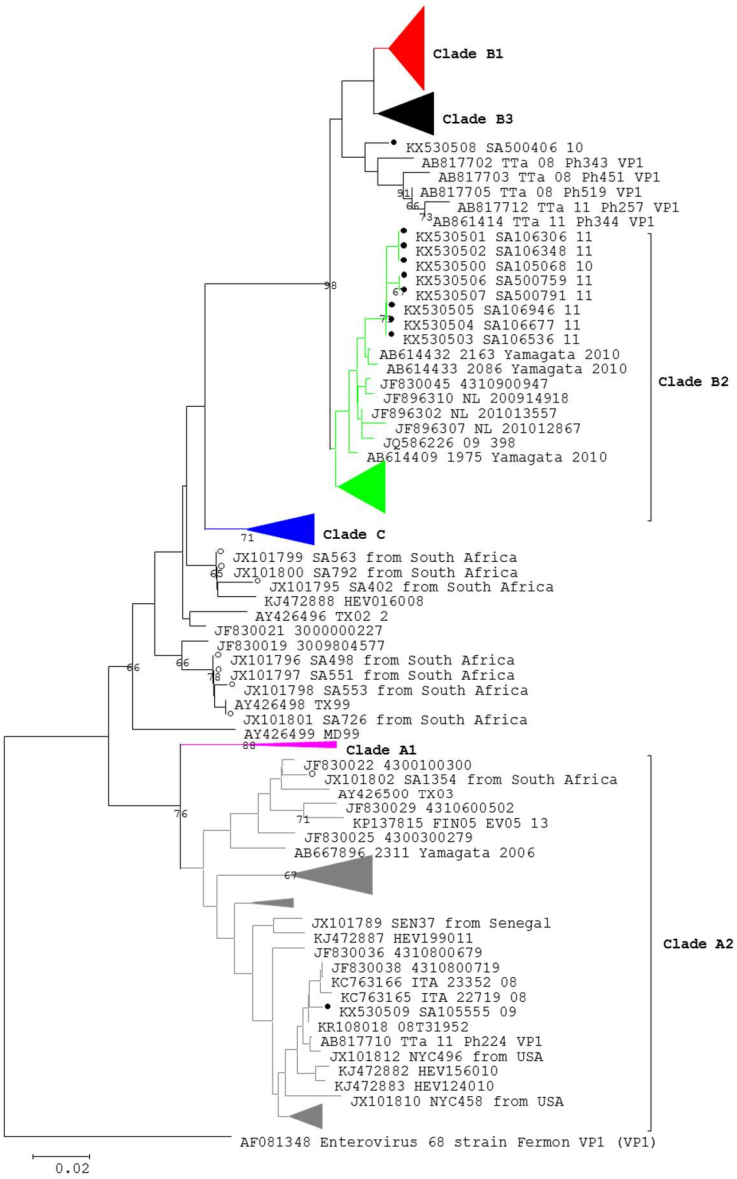


Figure 3.3: Molecular characterization of EV-D68 clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implicated in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Scale bar indicates number of nucleotide substitutions per site. Black dots = South African strains 2009-2011; Open dots = South African strains 2000-2001. Clades B1, B3, C A1 and partial B2 and A2 have been compressed to show comparison of South African strains to strains from other countries (country and year in blocks). Total number of sequences = 312 (representative strains from 2000-2016).

3.4 Discussion

During 2009-2011, enterovirus was detected in 8.2% of hospitalised SARI patients, which is within the range reported in other studies (3-25%) (39,78,79). Seventeen enterovirus genotypes were identified representing all four enterovirus species. EV-D68 was detected in 22% of enterovirus cases characterised (0.1%, 10/10260 amongst the total of hospitalised patients), which compares to frequencies reported in studies from The Netherlands (25%) and Germany (7.7%) (80,81). The EV-D68 strains that circulated in South Africa in 2009-2011 clustered in clades A and B of the EV-D species with the majority of strains grouping in clade B. All EV-D positive patients in this study were less than 5 years of age, which is similar to the typical younger age distribution reported for EV (78). A study to determine the worldwide emergence of EV-D68 included eight South African EV-D68 strains identified in hospitalised children over a one year period (2000-2001) (82). One of these South African strains (2001) was found within clade A, while the remaining seven (2000-2001) formed a cluster with sequences from Europe and the USA ancestral to clades B and C (82). More recent samples (2009-2011) were predominantly located in clade B of EV-D68.

It is speculated that the 2014 North American EV-D68 outbreak was likely due to modifications of the receptor interaction affecting the local cell tropism; the inability of previous established anti-EVD68 antibodies to detect the altered amino acid residues of the VP1gene (83).

A study from hospitalised patient with respiratory illness from New Zealand in 2010, indicates that EV-D68 may be associated with a more severe clinical presentation than other enterovirus strains (84). However, although the South African patients with EV-D68 detected in this study were all hospitalised, none were admitted to ICU or died. No recorded symptoms of neurological disease or paralysis for these patients was available, and no causative link between EV-D68 infection and neurological disease have been established to date (80). No EV-D positive samples were obtained from a South Africa study characterizing non-polio enteroviruses in stool and cerebrospinal fluid samples of patients symptomatic for gastroenteritis and acute

flaccid paralysis, respectively (personal communication, Wayne Howard, National Institute for Communicable Diseases). Higher EV-D68 activity in 2011 is consistent with another report (85). All polioviruses detected in this study were Sabin strains detected in the respiratory specimens of children that had recently been vaccinated.

This study has limitations that merit discussion. Only 24% of enterovirus positive samples selected for this study could be genotyped. The low success rate for genotyping may be due to low viral loads in clinical samples (Ct values >35 on original qRT-PCR assay) and subsequent inefficient nested amplification and sequencing of variable regions for phylogenetic analysis by degenerate PCR primers used for direct sequencing of PCR products. Furthermore, the study was limited to patients hospitalised with SARI and the presence of neurological manifestations was not specifically explored at time of participant enrolment. Identification of some enterovirus positives as rhinoviruses by molecular characterization could likely be due to cross reactivity in the 5` UTR due to sequence homology between enteroviruses and rhinoviruses.

Further investigation is needed to determine the disease association of enteroviruses and specifically EV-D68 infection with non-respiratory illness as well as full genome sequencing to determine possible recombination of EV-D68. Since this study did not include an asymptomatic control group, the study was limited in interpreting the attributing causality to enteroviruses in the presence of high frequency co-infections as well as determining concurrent prevalence in the healthy population.

In conclusion, the study showed that enteroviruses frequently contributed to SARI in South Africa. A high diversity in the enterovirus species that circulated in South Africa during 2009-2011, EV-D68 specifically circulated at high frequencies among children <5 years of age. Enteroviruses should be closely monitored to assess the emergence of highly pathogenic strains. There is no vaccine or antiviral available for non-polio enteroviruses. Understanding the contribution of enterovirus genotypes to severe illness will allow for informed decision making on potential candidate vaccines and development of therapeutic interventions.

Chapter 4

Enterovirus Genotypes among Patients with Severe Acute Respiratory Illness, Influenza-Like Illness and Asymptomatic Individuals in South Africa, 2012-2014

Reference:

Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, Madhi SA, du Plessis M, Cohen C, Treurnicht FK. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. Journal of Medical Virology 2017 Oct; 89(10):1759-1767

4.1 Introduction

Enteroviruses can cause symptoms similar to a mild cold, but have also been associated with severe respiratory infection requiring hospitalisation and may be fatal (86).

Though enteroviruses have been associated with SARI and ILI (21), these viruses have been less well characterized in South Africa. We identified high diversity among enteroviruses circulating in hospitalised South African patients (87), however it is unknown whether different enterovirus genotypes are associated with mild or severe respiratory illness in South Africa.

The aim was to characterize the enterovirus genotypes circulating among South African patients with SARI, ILI and asymptomatic individuals from June 2012 to

July 2014. In addition, the study assessed the association of different enterovirus species with mild (ILI) or severe (SARI) illness compared to asymptomatic individuals.

4.2 Materials and Methods

4.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI and ILI surveillance programmes as well as asymptomatic individuals from June 2012 through July 2014 (described in section 2.1).

4.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2 and 2.3.

Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (14) (methods section 2.4.1-2.4.2).

4.2.3 Amplification and sequencing of enteroviruses

The 5’ proximal part of VP1 region were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (described in section 2.4.3).

The genetic diversity of each VP1 sequence was first determined by comparison with the reference strains in GenBank (US National Center for Biotechnology

Information, NCBI, http://www.ncbi.nlm.nih.gov/, accessed 01 July 2016) using BLAST and confirmed by phylogenetic analysis of the partial VP1 sequences. Sub analysis was similarly done for the three dominant enterovirus species in this study. Sequences of enterovirus partial VP1 genes generated in this study have been deposited in GenBank with the following accession numbers: KX940982-KX941096.

4.2.4 Statistical analysis

The Chi-squared or Fisher's exact test were used for comparison of categorical variables. Unconditional exact logistic regression was used to assess the association of enterovirus species with disease severity among patients with mild (ILI) or severe illness (SARI) using asymptomatic individuals as control group. Exact logistic regression was used to account for the fact that no EV-D species was detected among controls. EV-B was used as the comparison group as it was the species most frequently detected.

Significance was assessed for $p<0.05$. All models were adjusted for age, HIV serostatus and underlying medical conditions. The analysis was performed using STATA 14.1 (StataCorp, College Station, Texas, USA).

4.3 Results

4.3.1 Epidemiology of the study population

During the study period 1525 patients with SARI, 3088 with ILI and 1401 controls were enrolled, of which 1494 (97.60%), 3079 (99.7%) and 1367 (97.6%) were tested and had enterovirus results available, respectively. The detection rate for enterovirus was 5.8% (87/1494), 3.4% (103/3079) and 3.4% (46/1367) among SARI, ILI and controls, respectively (SARI/controls, p=0.002 and ILI/control, p=0.973).
In children <1 year of age, the majority (61.1%, 44/72) of enterovirus positive samples presented with SARI while in the following age groups: 1-4 years (43.1%,

50/116), 5-18 years (53.9%, 14/26) and >18 years (81.8%, 18/22) the majority of patients presented with ILI ($p<0.001$). Among enterovirus positive samples stratified in age groups above, HIV infection was not significant in any of the cases or controls (data not shown). Enteroviruses were detected throughout the year among SARI/ILI cases and controls (Figure 4.1 A-C), with peaks above 10% during February – March, August and November.

4.3.2 Phylogenetic comparison of enterovirus genotypes identified during June 2012 – July 2014

Of the 236 enterovirus -positive specimens detected during the study period among SARI and ILI cases and controls, 30.1% (71/236) were excluded from VP1 RT-PCR attempts, since the enterovirus load in these samples was considered low (Ct-value >35) and the attempt has a higher probability of failure. An additional 16.5% (39/236) of samples were excluded from VP1 RT-PCR and sequencing attempts, because the original specimen volume was insufficient for RNA extraction.

Of the enterovirus -positive samples that were sequenced (53.4%, 126/236), 7.9% (10/126) could not be typed. Possible reasons for typing failure included failed sequencing reactions, mixed enterovirus infections (co-infections) and rhinovirus that was detected on the initial enterovirus screen qRT-PCR assay. No cloning was attempted to rule out mixed infections. Rhinovirus were identified in 11.9% (15/126) samples, and were excluded from the enterovirus phylogenetic analysis. Therefore, only 101 enterovirus -positive samples were used for subsequent phylogenetic comparison.

Neighbor-joining phylogenetic comparison of enteroviruses detected in clinical samples from June 2012 to July 2014 to international reference sequences (downloaded from GenBank) indicated that EV-B (60.4%; 61/101) and EV-A (21.8%; 22/101) were more commonly identified than EV-C (10.9%; 11/101) and EV-D (6.9%; 7/101) (Figure 4.2). The majority of EV-A (40.9%, 9/22) and EV-D

(57.1%, 4/7) were detected among ILI cases and the majority of EV-B (55.7%, 34/61) and EV-C (54.6%, 6/11) were detected among SARI cases (p=0.03).

On multivariable logistic regression analysis using enterovirus species B as the reference group and controlling for age, HIV status and underlying illness, no association with disease severity (SARI or ILI compared to controls) was identified for any of the enterovirus species (Table 4.1).

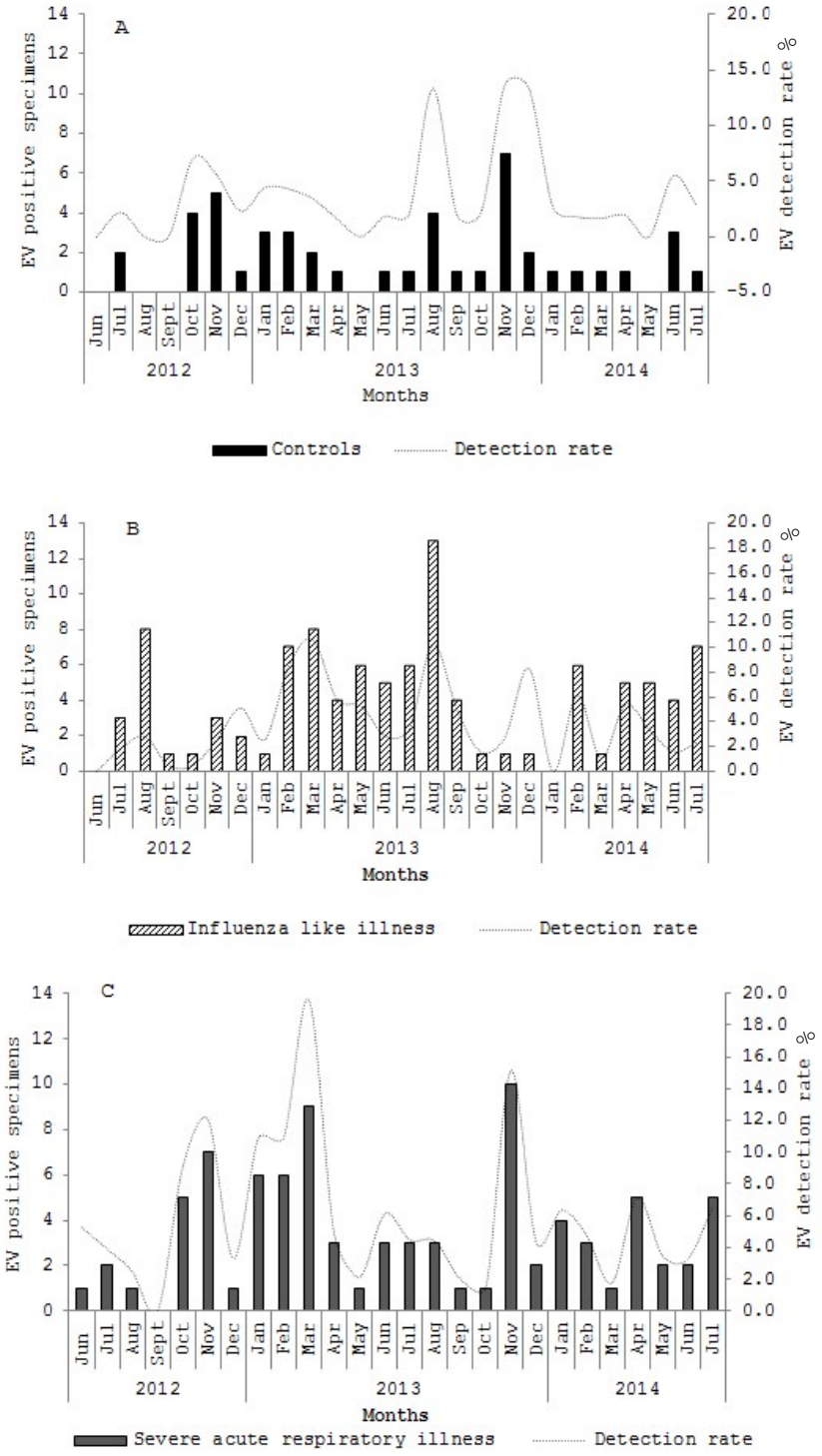


Figure 4.1: Monthly number and overall detection rate of enterovirus (EV)-positive cases among controls (A), influenza-like illness (B) and severe acute respiratory patients (C), Klerksdorp and Pietermaritzburg, South Africa, 2012-2014.

4.3.3 Enterovirus group and genotypes

Enterovirus genotypes could be identified in 59% (51/87), 33% (34/103) and 35% (16/46) of enterovirus-positive samples among SARI and ILI cases and controls, respectively. A total of 33 genotypes were identified, distributed among all four enterovirus species. The most prevalent genotype was E30 (9.9%, 10/101), followed by coxsackievirus B5 (7.9%, 8/101) and EV-D68 (6.9%, 7/101). All other genotypes identified in the study were detected in <5% of genotyped samples (Figure 4.3).

Coxsackievirus A14 (CVA14), echovirus 11 (E11), echovirus 16 (E16), echovirus 17 (E17), echovirus 19 (E19), echovirus 7 (E7), echovirus 9 (E9), poliovirus 1 (Sabin vaccine strain), poliovirus 3 (Sabin vaccine strain) were detected only among SARI cases and coxsackievirus A10 (CVA10), coxsackievirus A16 (CVA16), coxsackievirus B1 (CVB1), echovirus 32 (E32) and coxsackievirus A20 (CVA20) were detected only among ILI cases. Coxsackievirus B5, E30, echovirus 6 (E6) and EV-D68 were detected among SARI and ILI cases only (Figure 4.3).

EV-A, EV-B, EV-C and EV-D genotypes co-circulated in 2012-2014, while EV-D68 was detected sporadically in March-May in 2013 and 2014 only (Figure 4.4). The majority of enterovirus positive samples that could be genotyped, were from children aged <5 years (89.1%, 90/101) in the EV-B species (60.4%, 61/101) compared to the other enterovirus species, but were not statistically significant (p=0.08).

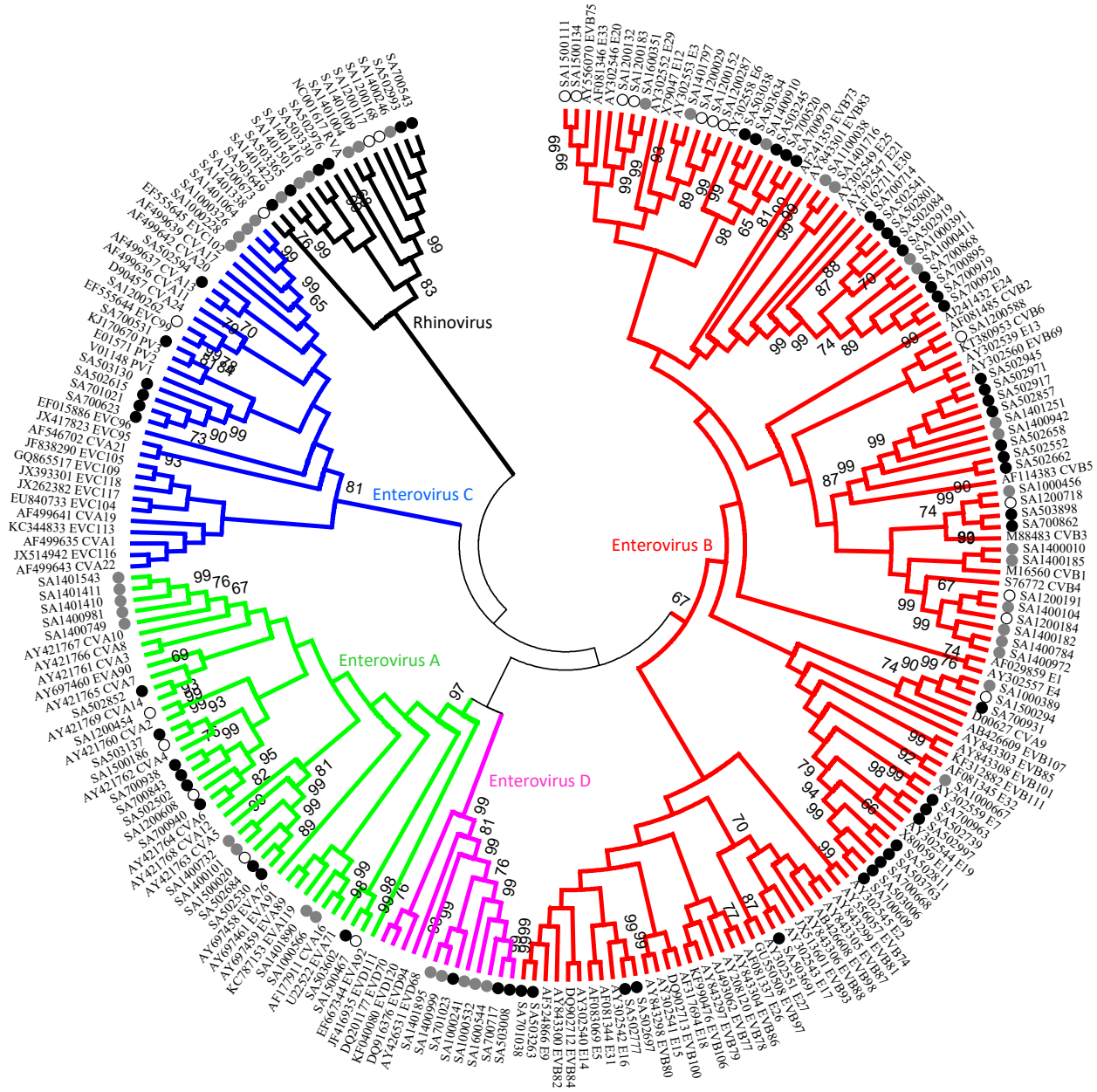


Figure 4.2: Phylogenetic analysis of enterovirus genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denotes genotypes identified in SARI patients, those with grey closed circles denotes genotypes identified in ILI patients while those with open circles denoted genotypes identifies in asymptomatic individuals. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree.

Table 4.1: Association of enterovirus species (EV-A to EV-D) with influenza-like illness (ILI) and severe acute respiratory illness (SARI) compared with asymptomatic controls, South Africa, 2009-2010 (bold indicate factors significant at p<0.05)

Factor	Prevalence			Multivariable analysis	
	Control	ILI	SARI	ILI	SARI
	n/N (%)	n/N (%)	n/N (%)	aOR* (95% CI)	aOR* (95% CI)
EV-A	5/16 (31)	9/34 (26)	8/51 (16)	1.16 (0.15-8.95)	1.43 (0.22-9.22)
EV-B	10/16 (63)	17/34 (50)	33/51 (66)	Reference	Reference
EV-C	1/16 (6)	4/34 (12)	6/51 (12)	0.41 (0.02-6.98)	0.68 (0.05-9.63)
EV-D	0 (0)	4/34 (12)	3/51 (6)	1.24 (0.11-∞)	0.53 (0.39-∞)

*Estimated odds ratio (aOR) adjusted for age, HIV and underlying medical condition

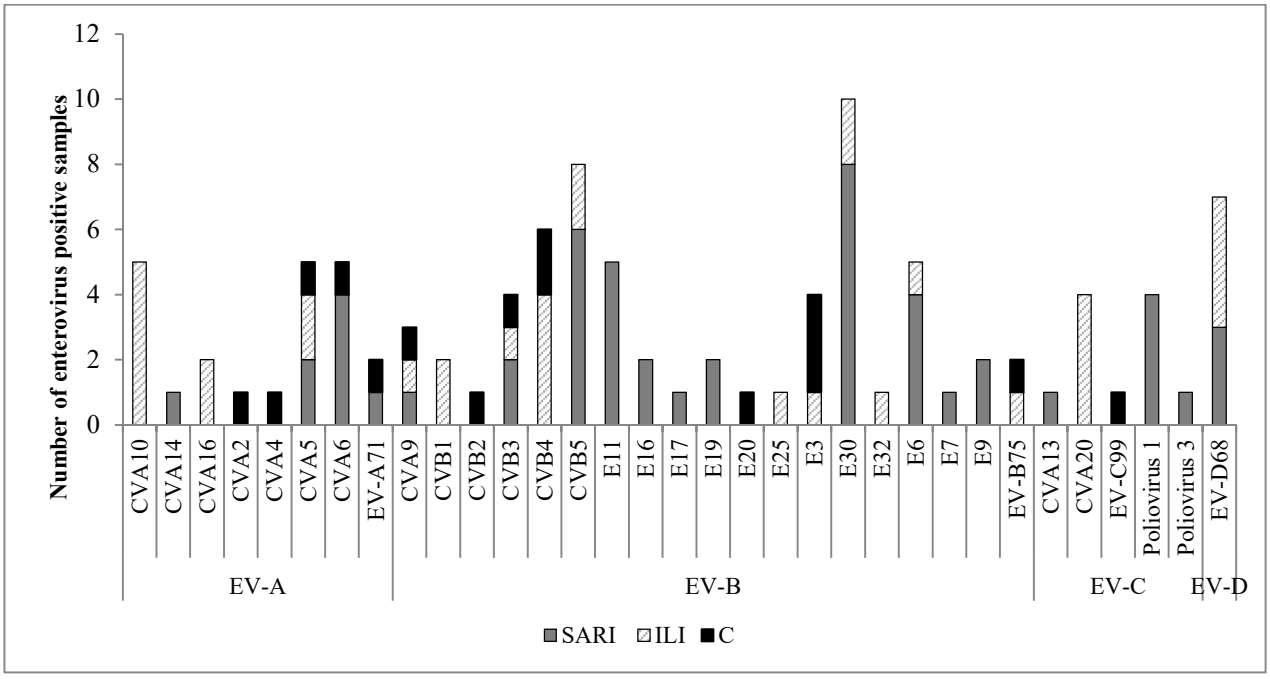


Figure 4.3: The genotypes and species (EV-A to EV-D) of enteroviruses identified in patients with severe acute respiratory illness (SARI), influenza-like illness (ILI) and asymptomatic individuals (C) from June 2012 to July 2014, South Africa

(CVA10=Coxsackie virus A10, CVA14=Coxsackie virus A14, CVA16=Coxsackie virus A16, CVA2=Coxsackie virus A2, CVA4=Coxsackie virus A4, CVA6=Coxsackie virus A6, EV-A71=Enterovirus A71, CVA9=Coxsackie virus A9, CVB1=Coxsackie virus B1, CVB2=Coxsackie virus B2, CVB3=Coxsackie virus B3, CVB4=Coxsackie virus B4, CVB5=Coxsackie virus B5, E11=Echovirus 11, E16=Echovirus 16, E17=Echovirus 17, E19=Echovirus 19, E20=Echovirus, E25=Echovirus 25, E3=Echovirus 3, E30=Echovirus 30, E32=Echovirus 32, E6=Echovirus 6, E7=Echovirus 7, E9=Echovirus 9, EV-B75, CVA13=Coxsackie virus A13, CVA20=Coxsackie virus A20, EV-C99=Enterovirus C99, EV-D68=Enterovirus D68)

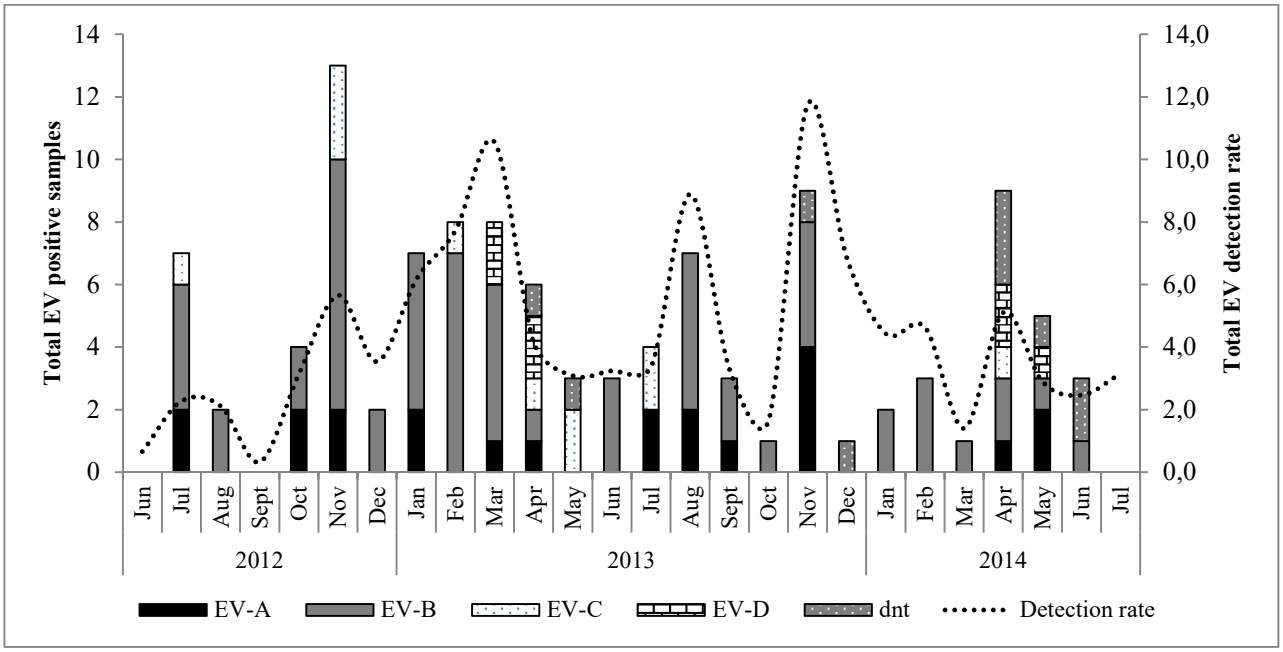


Figure 4.4: Distribution of enterovirus (EV) species (A-D) in South Africa during June 2012 to July 2014; dnt=did not genotype

4.3.4 Molecular evolution of echovirus 30, coxsackie virus B5 and EV-D68 genotypes

4.3.4.1 Echovirus 30

Phylogenetic analysis of E30 samples showed that all E30 samples from this study clustered with a single Australian strain from 2000 (GU232849) with 98% bootstrap support. Following the classification convention for E30 as described by Bailly *et al* (2009), this study propose to call this sub cluster genotype k (2). The sequences depicted in Figure 4.5 represent all known genotypes.

All E30 strains identified in this study displayed >98% nt homology and formed a distinct bi-phyletic cluster with 79% bootstrap support (Figure 5). In the bi-phyletic cluster, strains from Edendale SARI cases and Klerksdorp SARI and ILI cases are distinctly located in each of the sub clusters with 92% and 93% bootstrap support, respectively.
The E30 strains in this study differ from other E30 genotypes by mean nucleotide pairwise distance of 20.2% (15.6-27.5) and mean amino acid sequence distance of 6% (2.7-11.6) in amino acid sequence.

4.3.4.2 Coxsackie virus B5

All coxsackie virus B5 strains identified in this study (6 SARI samples and 2 ILI samples from Edendale) were categorized as genogroup C and formed a distinct sub cluster with a bootstrap value of 91% (Figure 4.6). The nucleotide sequences for coxsackie virus B5 strains were downloaded from GenBank and sequences depicted in Figure 6 represent all known genogroups. 6/8 samples (4 SARI and 2 ILI samples) were identical (nt homology 100%). All of the coxsackie virus B5 strains (identified in this study) displayed close genetic relationships with samples identified in surface water from Belarus, 2007 (GQ352398) and faeces from Germany, 2009 (HF948277).

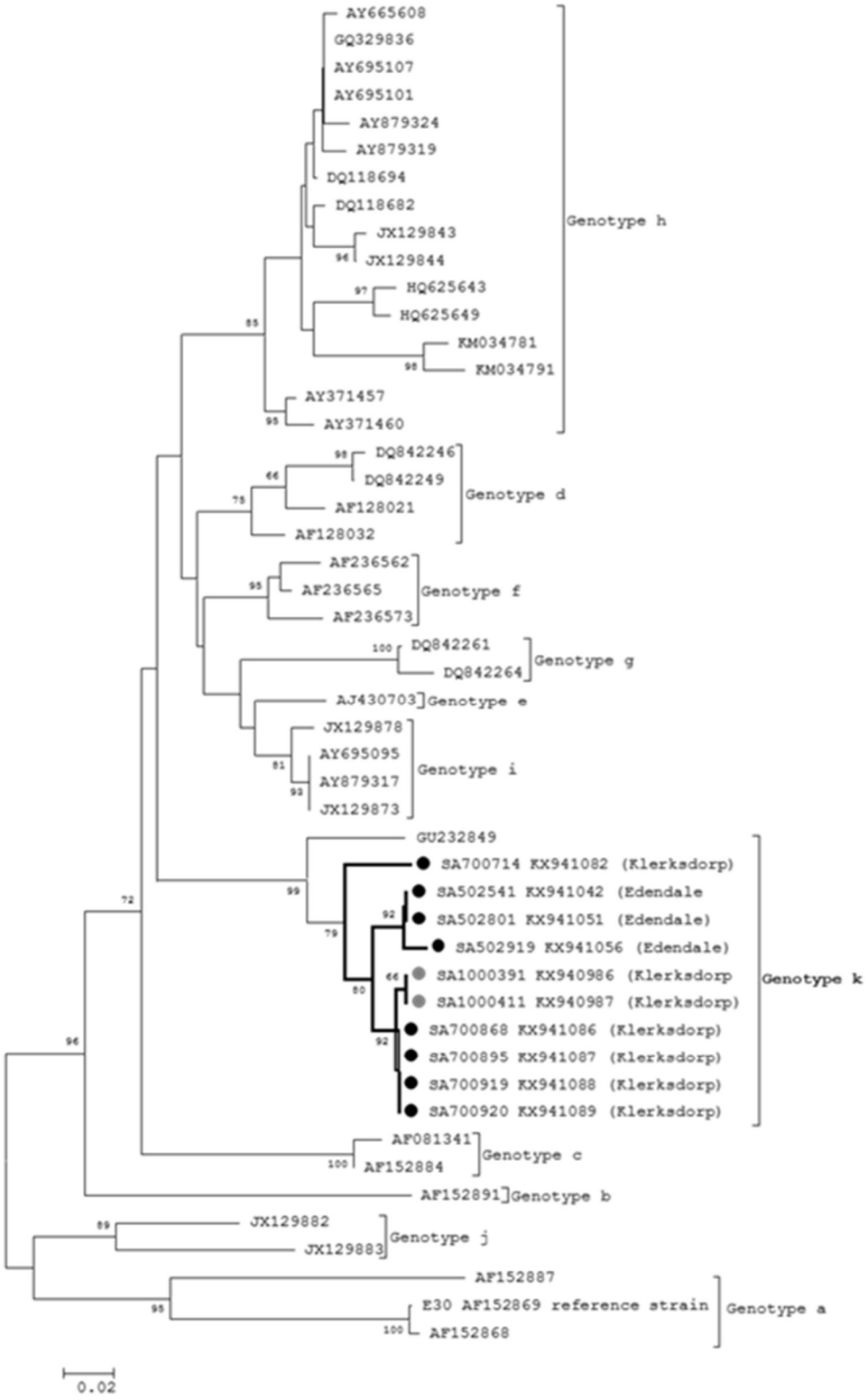


Figure 4.5: Phylogenetic analysis of echovirus 30 genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

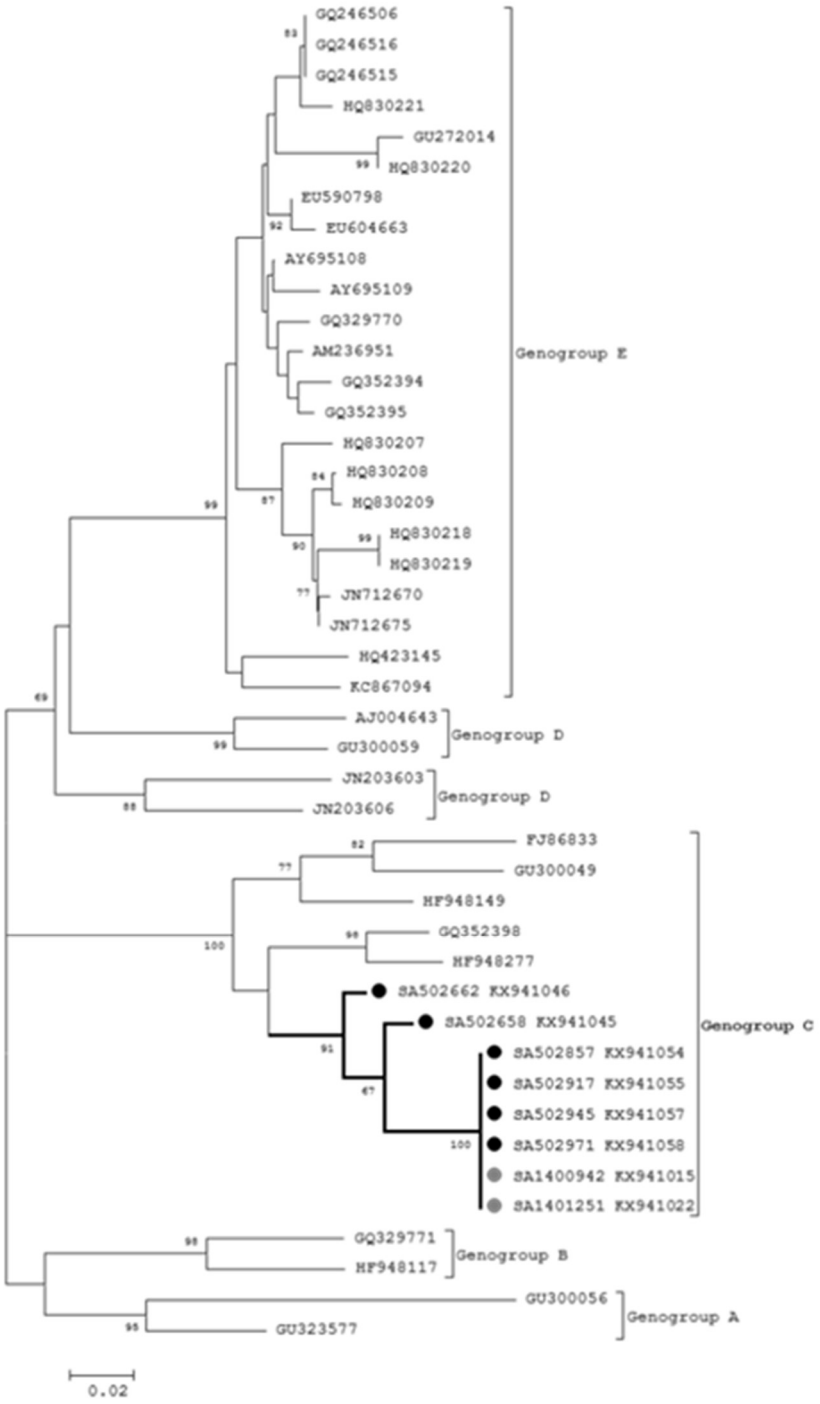


Figure 4.6: Phylogenetic analysis of coxsackie virus B5 genogroups by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

Sequences from the South African coxsackie virus B5 strains differed from the other coxsackie virus B5 genogroup C sequences by 13.4% in pairwise nucleotide sequence distance and 3% in pairwise amino acid sequence. Strains of coxsackie virus B in samples 502662 and 502658 which were not identical to the other South African sequences were sampled in 2012 and 2013, respectively.

4.3.4.3 EV-D68

The majority (71.4%, 5/7) of EV-D68 strains identified in 2013/2014 distinctly clustered in lineage B2 (89% bootstrap support) of which 3 were ILI cases and 2 were SARI cases, together with strains from one of the two co-circulating EV-D68 lineages that caused the large 2014 USA outbreak (Figure 4.7). The nucleotide sequences for EVD-68 strains were downloaded from GenBank and sequences depicted in Figure 4.7 represent all known lineages.
The lineage B EV-D68 strains in this study differed from other lineage B EV-D68 strains by 1.5% nucleotide and 1.4% amino acid sequence. These included 2 SARI cases (1 from Klerksdorp and 1 from Edendale) and 3 ILI cases (1 from Klerksdorp and 2 from Edendale).

The other South African EV-D68 strains (28.6%, 2/7) belonged to lineage A forming a unique cluster. The two strains from 2013 are from the same geographic location and clustered with 99% bootstrap support. The lineage A EV-D68 strains in this study differed from other lineage A EV-D68 strains by 12.8% nucleotide and 3.0% amino acid sequence. The 2012-2014 South African EV-D68 strains did not cluster with strains of EV-D68 previously identified in South African patients ((34) ,unpublished data).

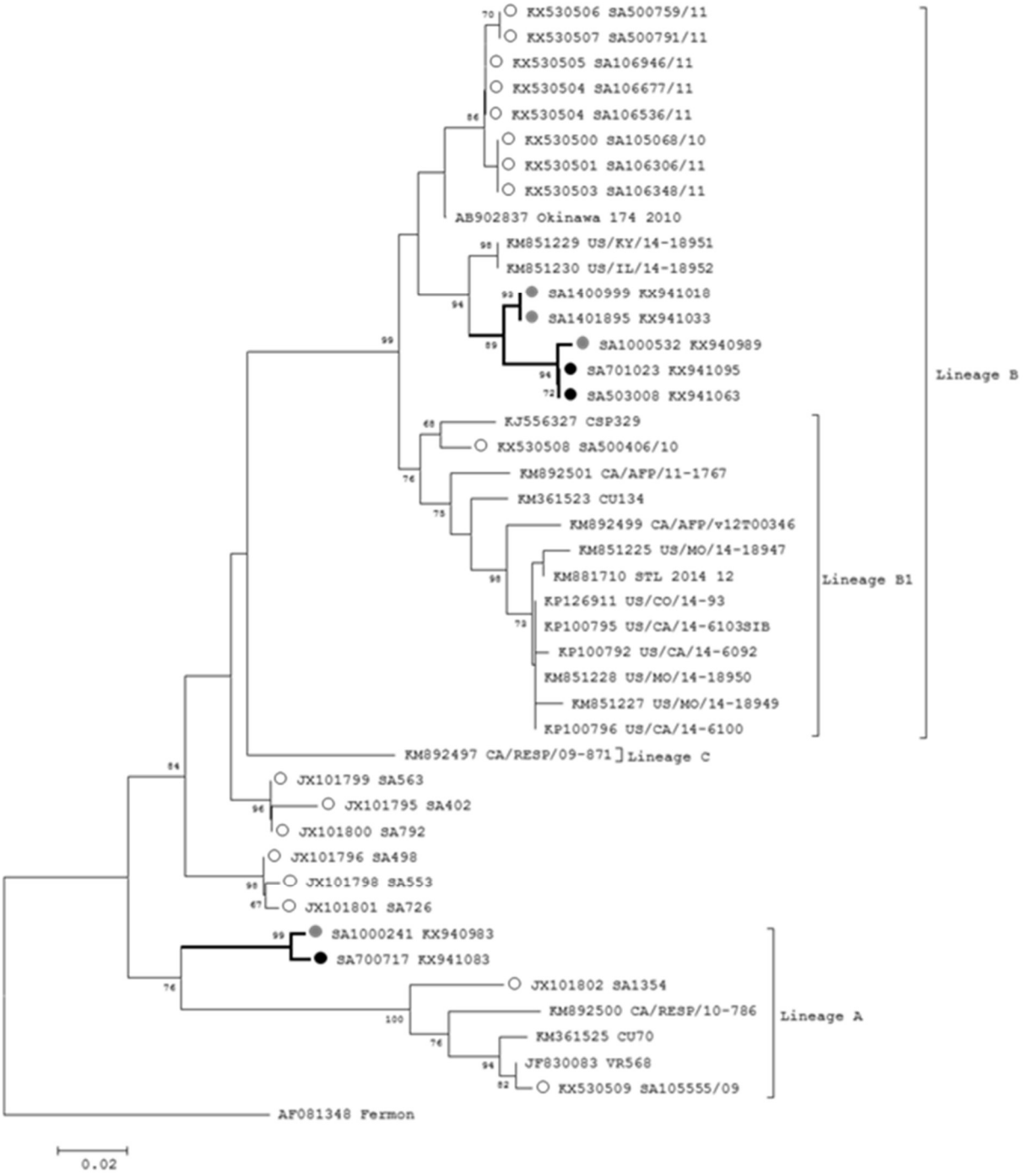


Figure 4.7: Phylogenetic analysis of EV-D68 lineages by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Sequences with open circles are EV-D68 strains previously identified from South Africa. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

4.4 Discussion

This study observed a high diversity of circulating enterovirus genotypes (a total of 33 genotypes) from enterovirus species A-D with high circulation rates for EV-B and EV-A compared to EV-C and EV-D. No disease association of enterovirus species with disease severity were determined.

From the systematic surveillance across all age groups, the majority of patients testing positive for enterovirus were < 5 years of age. This study did not observe any difference in disease association due to different enterovirus species. This could be firstly because of a lack of power (small sample numbers) and since enterovirus disease association cannot be conclusively ascribed since the analysis did not include co-infections with other pathogens. Epidemics of human enterovirus disease display a seasonal pattern, with infections more common in summer and early autumn in geographical regions with a temperate climate (3,38). Most clusters of EV-D68 showed an atypical late seasonality compared to other enteroviruses, with a peak in autumn, instead of summer (3). Outbreaks of enterovirus disease in autumn, early or late, are typical. Distribution of EV-A through –D in most temperate zones is from spring through autumn, with peak disease most years in late summer-autumn. This study showed similarities in the seasonal distribution of the enterovirus species detected in South Africa as enterovirus species A-D were detected throughout the year, but EV-D68 was sporadically detected in autumn (March-May) during 2013-2014.

E30, coxsackie virus B5 and EV-D68 were the genotypes most frequently identified in this study. Within a genotype, enteroviruses are deemed relatively uniform, as was observed in this study for E30 and coxsackie virus B5. However, some studies have suggested variation within a genotype (88), such as was observed for EV-D68 in this study.

E30 is the most common enterovirus strain responsible for sporadic as well as large-scale outbreaks of aseptic meningitis in many countries of Europe and North America (2). E30 causes large outbreaks of aseptic meningitis that attracts public

health attention and as a result there is a rich and large database of E30 VP1 sequences regarding aseptic meningitis, E30 has however not been widely associated with respiratory illness.

A common feature of E30 molecular epidemiology is the progression of circulating lineages within one prevalent genotype (2,18). The E30 strains identified in this study clustered together (designated genotype k) although no differences were observed between viruses from SARI compared to ILI cases. There is no report of a concurrent E30 aseptic meningitis outbreak at the same time the E30 combined nasopharyngeal and oropharyngeal specimen (NP/OP) positives were detected. It has been reported that genotypes of E30 can become dominant for 3-4-year periods before they disappear from circulation (88), and they are therefore known as epidemic strains. In this study, it seems that E30 genotype k became dominant and circulated in South Africa during 2012-2014. Phylogenetic studies with VP1 sequences indicate that E30 variants are continuously emerging and replacing the circulating E30 strains, largely due to the error prone enterovirus RNA polymerase (2).

Coxsackie virus B5 has been detected for over 50 years and, similar to E30 , sporadic cases of aseptic meningitis, as well as outbreaks, have been reported, remaining one of the most predominant reported enterovirus genotypes in a number of countries (89). All of the coxsackie virus B5 strains identified in this study formed a sub cluster in genogroup C. Phylogenetic clustering by year of study enrolment was observed for the South African strains. Different co-circulating strains of enterovirus are often observed in enterovirus disease outbreaks as enterovirus genomes vary with respect to time but may also vary by geographical distribution (90).

In 2014, the largest outbreak to date of severe respiratory illness associated with EV-D68 occurred in the USA, and almost all of the case patients were children (91). Coinciding with the 2014 USA outbreak there was an increased number of reports of children with sudden onset acute flaccid myelitis. Some recent South

African EV-D68 strains clustered with USA 2014 lineage B2 EV-D68 strains. The USA 2014 lineage B2 co-circulated with lineage B1 EV-D68 during the large 2014 USA outbreak. EV-D68 lineages B1 and B2 are both temporally associated with acute flaccid myelitis (92). The majority of EV-D68 strains identified in this study clustered in lineage B2 and the remaining samples clustered in lineage A, forming a unique cluster. The 2012-2014 South African samples did not cluster with strains previously identified in South Africa during 2009-2011.

The mechanisms for emergence of enterovirus-associated outbreaks are not known; however, a combination of virus-specific, population-level and other external factors are likely to be involved (3).

This study has some limitations. Previous studies observed a difference in the predominant enterovirus genotypes among neonates and older children (93) but this study was not powered for this analysis. Some rhinovirus sequences were detected in samples that tested positive for enterovirus using the in-house assay indicating that this enterovirus detection assay is not 100% specific to detect only enteroviruses. This study was not powered to describe the role of co-infections of other respiratory pathogens, and there for the role of enterovirus causing SARI and ILI could be overestimated.

In conclusion, this study showed that there was a high diversity in the VP1 sequences of the enterovirus species circulating in South Africa during 2012 to 2014 and most of these genotypes are associated with meningitis worldwide. Enterovirus was detected in outpatient and hospitalised patients with SARI but was also detected in controls. No disease association of enterovirus species with disease severity were found; however, some genotypes (E30, coxsackie virus B5, EV-D68) were more prevalent in symptomatic cases. Further studies are needed to determine if other factors such as viral load or host interactions play a role in enterovirus -associated disease as well as a robust spatiotemporal phylogenetic analysis based on complete VP1 sequences.

Chapter 5

Genome Analysis of Human Echovirus 30 Isolated from Patients with Acute Respiratory Illness in South Africa, 2013

5.1. Introduction

E30 was the most frequently identified enterovirus genotype in symptomatic cases (overall 9.9%, 10/101) in a previous study from South Africa, which compared the enterovirus genotypes in a small group of patients with severe and mild respiratory illness to asymptomatic individuals (94). Other enteroviruses, like EV-D68, have putatively been associated with respiratory illness and neurological symptoms (95). Enteroviruses are also the most common cause of aseptic viral meningitis, an inflammation of the tissue that covers the brain and the spinal cord. Although aseptic meningitis is mostly self-limiting, young children and immune-compromised individuals are more likely to have severe illness (1). One enterovirus genotype, E30, causes large outbreaks of aseptic meningitis, which is of public health concern (2,3). Therefore, it is important to identify and describe the genotypes causing illness or emerging genotypes. As recombination in enteroviruses is well documented there is an increasing need to produce full-length viral genome sequences (Figure 5.1). The ORF are flanked on either side by UTR's and the P1 genomic region encodes four capsid or structural proteins (VP1, VP2, VP3 and VP4) while the P2 and P3 genomic regions encode seven non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C and 3D).

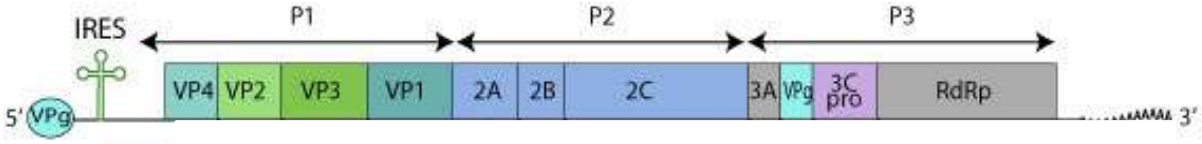


Figure 5.1: Genome structure of echovirus 30 (permission from ViralZone) (VPg – Viral protein genome linked; IRES – internal ribosomal entry site; RdRp – RNA dependent RNA polymerase)

Only nineteen E30 full length genome sequences are currently available in the public domain, mainly isolated from individuals with aseptic meningitis and only two isolated from individuals with respiratory illness in Thailand in 2010 (96). No full length E30 genome sequences are available from sub-Saharan Africa, or any countries with high HIV disease burden. E30 infections in HIV-infected individuals could lead to severe disease. Genome sequencing is important to identify evolutionary changes across the genome including potential inter-typic and intra-typic recombination events which might lead to emergence of contemporary strains with altered immune phenotypes or re-emergence of historic strains which may cause disease outbreaks (63,97).

Enrichment of samples for genome sequencing using next-generation sequencing can include RNA reverse transcription and cDNA amplification with random hexamers or alternatively with sequence independent single primer amplification (SISPA) primers (98,99). In some cases, the clinical samples can be filtered, treated with DNAse or subjected to a depletion of rRNA while in other cases combinations of these treatments can be used. The SISPA methodology for generation of whole genome sequences has not been described for enteroviruses before.

The aim of this study was to characterise E30 strains from respiratory illness cases across the complete viral genomes in South Africa from 2012 to 2014 using a SISPA-based next generation sequencing approach.

5.2 Materials and Methods

5.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI and ILI surveillance programmes as well as asymptomatic individuals from June 2012 through July 2014 (described in section 2.1).

5.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2.1. Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (14) (methods section 2.2-2.3).

5.2.3 Amplification and sequencing of enteroviruses

The 5' proximal part of the VP1 region was amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (Section 2.4.1-2.4.3)(100).

5.2.4 Enterovirus isolation and genome sequencing

E30 positive samples were inoculated into human epithelial type 2 (HEp-2) cells for virus isolation and passage of cultures as described in section 2.4.4.

Rapid amplification of cDNA ends followed by sequence independent single primer amplification (RACE-SISPA) of E30 genomes from propagated virus were done with and without enrichment for viral RNA (section 2.4.5-2.4.6). Nineteen

reference E30 genome sequences downloaded are annotated with a GenBank accession number followed by place and year of isolation. GARD (https://www.datamonkey.org/) examined 10664 models, at a rate of 0.49 models per second to identify recombination breakpoints between the South African strains and other E30 genogroups (n=24). The alignment contained 2517 potential breakpoints, translating into a search space of 3.96364736471951e+22 models. Comparing the Akaike information criterion with a correction for small sample sizes (AICc) score of the best fitting GARD model, suggests that because the multiple tree model can be preferred over the single tree model by an evidence ratio of 100 or greater, at least of one of the breakpoints reflects a true topological incongruence. Usually the best model is chosen to be the model with the smallest AIC while the delta AIC calculates the difference between the AIC of each model (76).

5.3 Results

5.3.1 Enterovirus isolation

Of the genotyped samples, 9.9% (10/101) were previously identified as E30 by partial sequencing (Chapter 4). Only 8 of the 10 E30 positive samples had residual volume left for virus isolation and 5 (62.5%) were successfully propagated. E30 cultures presented with classic enterovirus cytopathic effect in Hep-2 cell monolayers (characterised by rounded, retractile cells detaching from the substrate surface with cytoplasmic stranding). Infection was confirmed by RT-PCR for enterovirus as described previously (100). All cultured samples were positive with Ct values less than 20.

5.3.2 Characteristics of the echovirus 30 positive cases

Of all 10 patients positive for E30, 20% (2/10) were enrolled with ILI and 80% (8/10) were enrolled with SARI. The majority (80%, 8/10) were paediatric patients less than 2 year of age.

All five successful E30 isolates were from children <2 years of age [mean 8.0 months (range, 1.4-18.4 months)]. Four of the five E30 isolates were derived from hospitalised patients who met the SARI case definition and were enrolled in 2013 at two surveillance sites; Edendale (25%, 1/4) and Klerksdorp (75%, 3/4). The fifth isolate was from a patient who met the ILI case definition and also enrolled in 2013 from an outpatient clinic in Klerksdorp. All five patients were HIV uninfected, but one hospitalised patient was HIV-exposed. The median duration of illness was 1 day before admission. The median duration of hospital stay was 4 days (range 3-37). Fever or history of fever was reported for all patients. None of the patients had any underlying illnesses. All patients were discharged and had an uneventful recovery (100).

5.3.3 Genome sequencing

Total nucleic acid samples enriched for RNA only before double stranded cDNA preparation

Only 1 (sample 1000411) of the TNA extracts from 5 E30 cultures had sufficient template concentration and acceptable quality (smear between 300-100 bp) for NGS following enrichment for viral RNA (Figure 5.2).

Non-enriched TNA samples for double stranded cDNA preparation

The non-enriched TNA samples' double stranded cDNA concentrations were as expected significantly higher ($p<0.001$) than those obtained for the RNA-enriched

samples (Table 5.1). All five samples were processed for NGS library preparation sequencing and reads obtained ranged from 2353047 to 3510297 (Table 5.1).

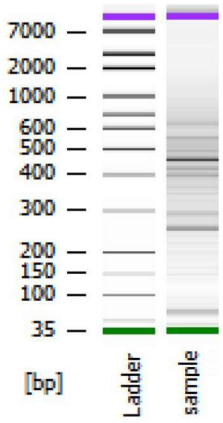


Figure 5.2: Example of electrophoresis of an echovirus 30 positive sample run on the Bioanalyzer, indicating smear between 300-1000bp

Table 5.1: Concentration of double stranded cDNA generated using RACE-SISPA comparing enriched and non-enriched viral samples as well as total reads generated with MiSeq

Sample ID	*TNA enriched for viral RNA [ng/µL]	**Non-enriched TNA [ng/µL]	Number of reads generated
502801	1.39	41.00	2481174
700714	1.55	33.60	2419220
1000411	3.73	110.30	2580378
700895	2.24	37.40	3510297
700919	1.61	20.00	2353047

*template for cDNA consists of RNA only; **template for cDNA consists of RNA and DNA

5.3.4 Complete genome analysis of echovirus 30 strains

The complete whole genomes were recovered by using reference mapping with the E30 prototype, Bastianni strain (AF311938) (Figure 5.3) after pairing and trimming of the total reads obtained for each sample as indicated in Table 5.1. The complete genome sequence consisting of 7453 nucleotides, of which 6582 nucleotides in an open reading frame encoding a polypeptide of 2194 amino acids

were recovered from the 5 E30 strains from South Africa. The coding sequences were flanked by a non-coding 5' UTR of 749 nucleotides and a non-coding 3' UTR of 99 nucleotides followed by a poly (A) tail composed of a long sequence of adenine nucleotides. Alignment of the full-length genome sequences from E30 isolates with the genome of the E30 Bastianni strain showed that they all had the same genomic organization and co-linear order of genomic regions.

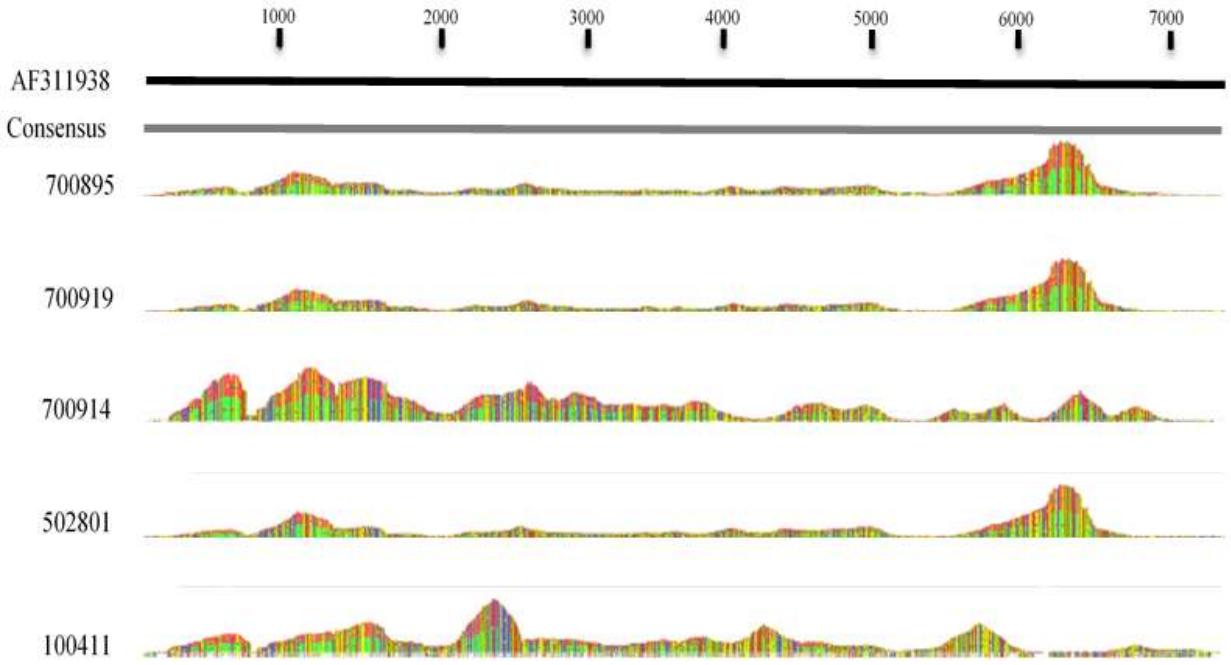


Figure 5.3: Schematic representation of genome coverage after reference mapping against the E30 prototype, Bastianni strain (AF311938)

Compared to the 19 E30 reference strains, the nucleotide sequences of the non-structural regions (P2 and P3) were more variable (74.4%-82.00% and 70.8%-85.1%, respectively), while the sequences of structural regions (P1) were more conserved (79.2%-85.4%) (Figure 5.4). This translated into amino acid changes and non-structural regions (P2 and P3) were more variable (76.5%-83.30% and 73.2%-86.2%, respectively), compared to the structural regions (P1) which were

more conserved (81.1%-86.4%). The complete genomes identified in this study have been deposited in GenBank (accession numbers MK376268-MK376272).

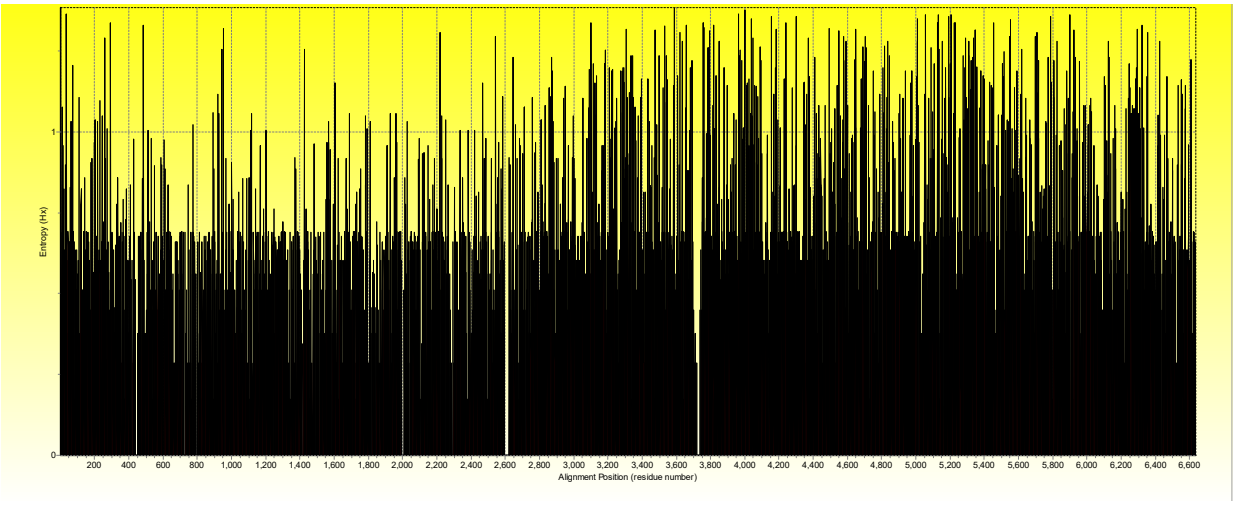


Figure 5.4: Entropy plot indicating the difference of the South African echovirus 30 strains across the polyprotein

5.3.5 Phylogenetic analysis

5.3.5.1 Phylogenetic analysis of complete polyprotein

Phylogenetic analysis of the polyprotein nucleotide sequences (6580 bp) displays a bi-phyletic clade supported by bootstraps of 72% and 94% respectively, for the South African E30 strains from 2013 (Figure 5.5). The South African viruses clustered with 94% bootstrap support with strains from the Netherlands, China and USA; where all South African strains are in a bi-phyletic sub-cluster supported by 100% bootstrap value. The complete genomes of South African E30 strains were also analysed with the online genotyping tool (https://www.rivm.nl/mpf/typingtool/enterovirus/) which could not assign a genotype.

Phylogenetic analysis were also done to assess if phylogenetic clustering observed for the complete polyprotein were maintained in the sub-genomic regions.

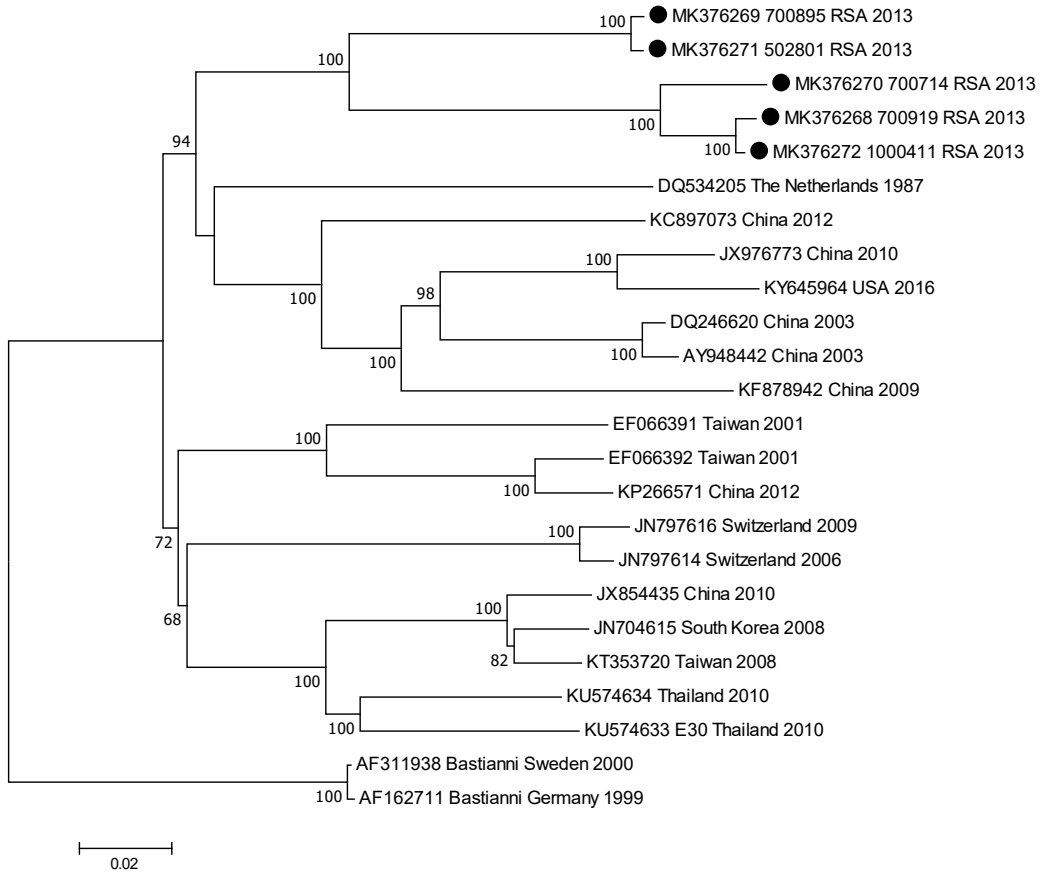


Figure 5.5: Genomic (6580 bp) relationship of 5 South African echovirus 30 (E30) strains compared to 19 other E30 strains available in GenBank database. Genome sequences were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstraps replicated). The scale bar represents the genetic distance. Bootstrap values >65% are shown. ● South African strains from this study.

5.3.5.2 Phylogenetic analysis of sub-genomic P1 region

The phylogenetic tree representing the P1 region (2580 bp) is indicated in Figure 5.6. The E30 strains from this study still clusters together with 100% bootstrap support.

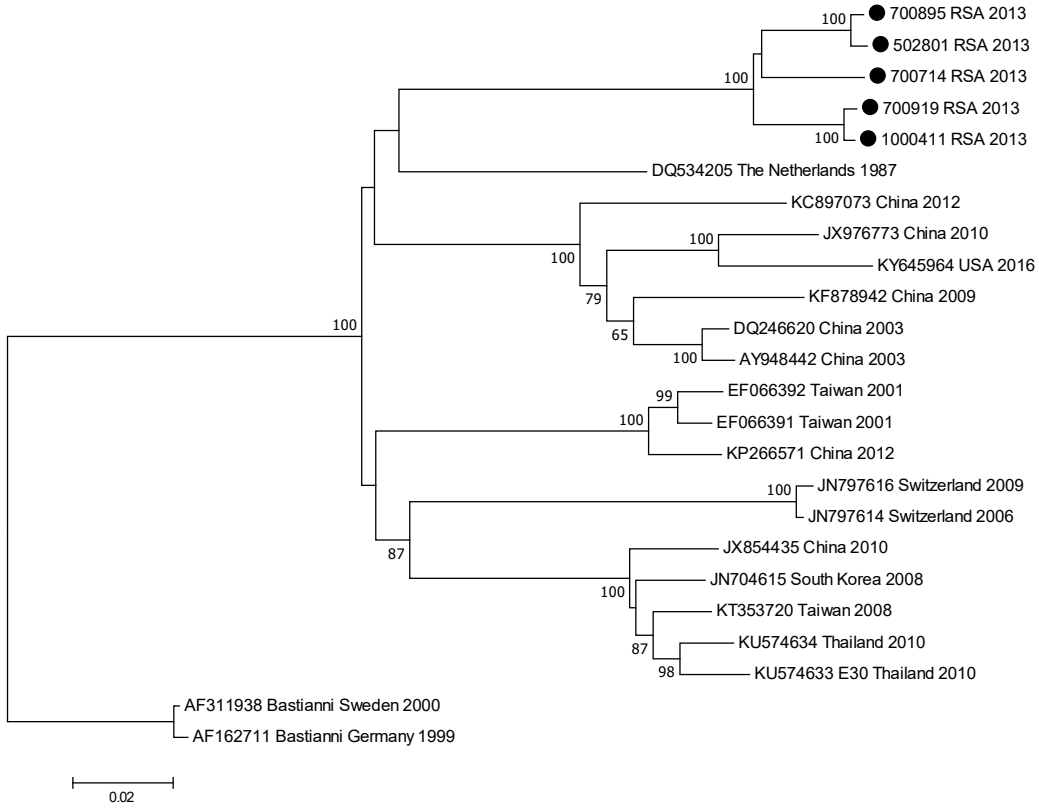


Figure 5.6: Phylogenetic relationships based on the complete P1 sub genomic region (2580 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

5.3.5.3 Phylogenetic analysis of sub-genomic P2 and P3 regions

The phylogenetic tree representing the P2 region is indicated in Figure 5.7. The South African strains did not form a single cluster, and two of the strains (MK376269 and MK376271) cluster with 70% bootstrap support with Asian strains from 2008 and 2010; including two strains (KU574633 and KU574634) from respiratory cases identified in Thailand in 2010. The other three South African strains form a separate sub cluster with strains from the Netherlands, China and USA.

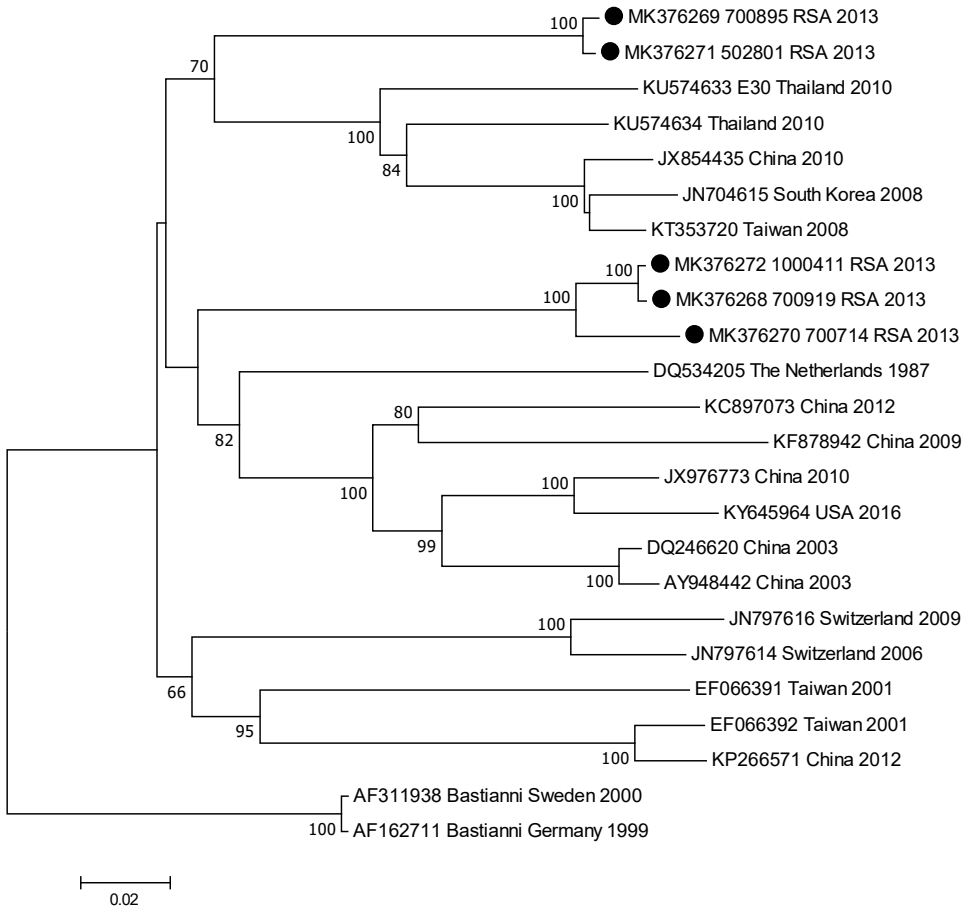


Figure 5.7: Phylogenetic relationships based on the complete P2 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

The phylogenetic tree representing the P3 region (2268 bp) is indicated in figure 5.8 and display similar clustering patterns for the South African E30 strains as was observed for the P2 region and the 2 subclusters are supported by 88% (MK376269 and MK376271) and 100% (MK376272, MK376268 and MK376270) bootstrap values respectively.

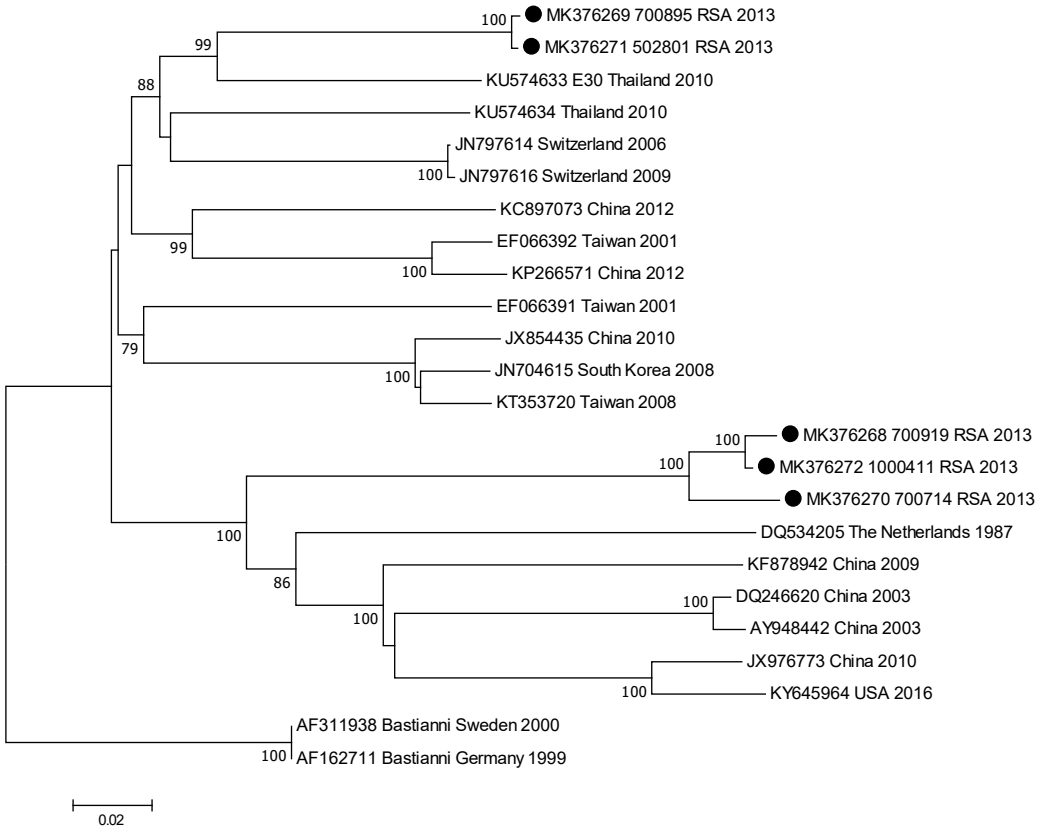


Figure 5.8: Phylogenetic relationships based on the complete P3 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

5.3.5.4 VP1 sub-genomic analysis

The phylogenetic tree representing the complete VP1 region (876 bp) is indicated in Figure 5.9. The five South African E30 strains still clustered separately (with 100% bootstrap support) to other E30 genogroups in a previously proposed genotype k.

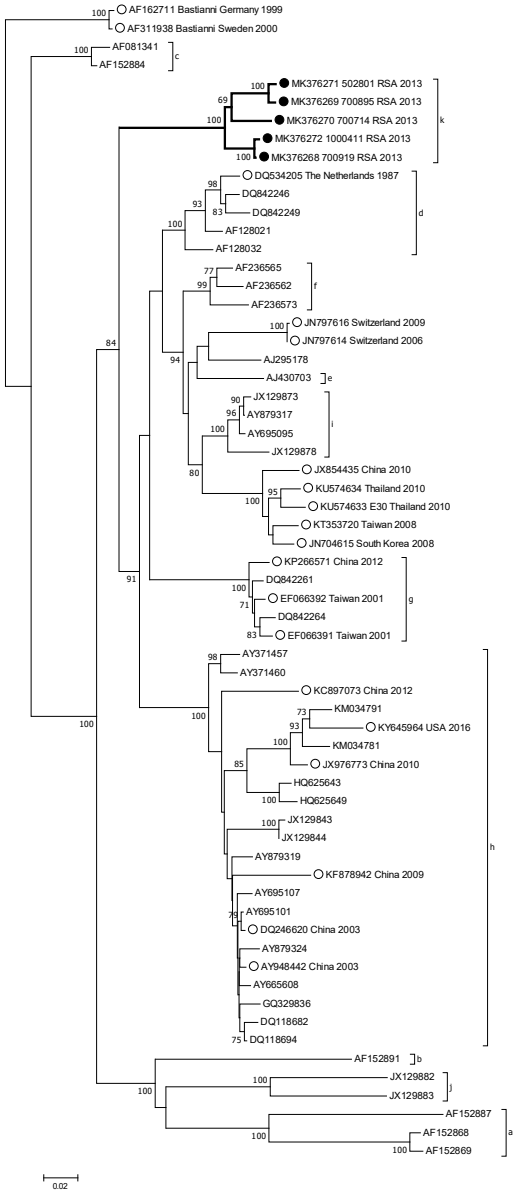


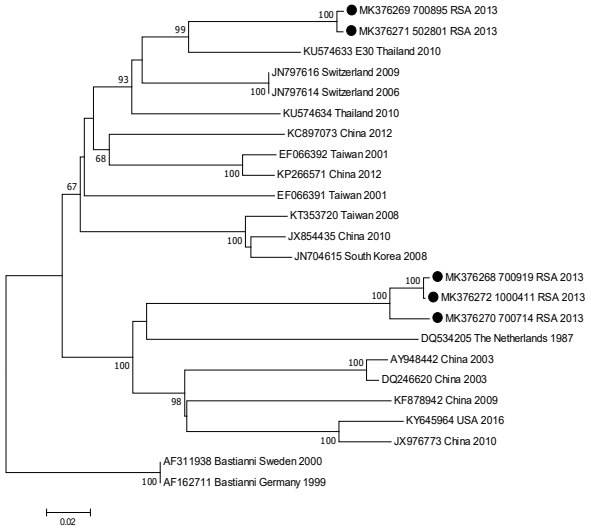
Figure 5.9: Phylogenetic relationships based on the complete VP1 genomic region (876 bp) of echovirus 30 (E30) strains. Five South African and 58 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis. Bold line indicates the proposed genotype k.

5.3.6 Investigation of E30 genomes for recombination

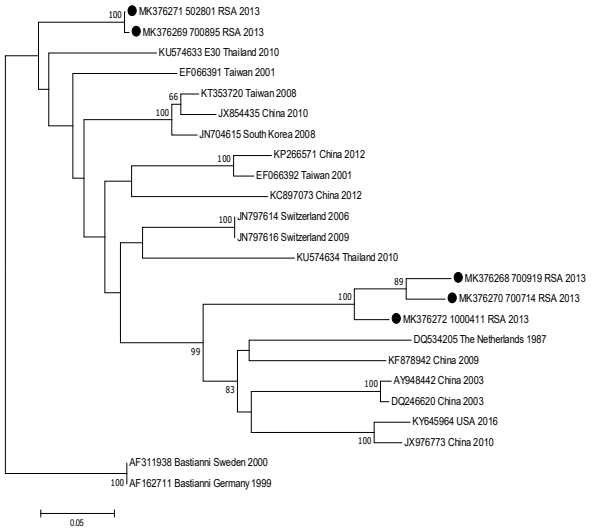
GARD identified at least 8 putative breakpoints of which breakpoints at nucleotide positions 2886, 4347, 5046 and 5364 were statistically significant (p= 0.0014, two-tailed; Figure 5.10). The phylogenetic trees for the various partitions at breakpoints confirms two distinct clusters in genomic regions outside of P1 and VP1. One cluster included two South African strains, MK376269 and MK376271 together with 2 European strains from 2006 and 2009 and 9 Asian strains ranging from 2001-2010. The other cluster consisted of three South African strains, MK376268, MK376272 and MK376270 together with 1 European strain from the 1980s, 4 Asian strains ranging from 2003-2010 and one recent USA strain (Figure 5.11). The clustering pattern of the South African strains did not change between the 4 partitions created by the significant breakpoints identified and confirms the distinct evolutionary history of the clusters. Further inter-typic recombination analysis should be done to identify the most recent common ancestor for the 2 clades in addition to generating more E30 genomes from various geographic regions and anatomical sites. Genomes were submitted to GenBank under accession numbers MK376268 to MK376272.

Topological incongruence report			
Breakpoints	LHS p-value	RHS p-value	Significance
2569	0.3654	0.0154	
2886	0.0014	0.0014	●●●●●●●●
3421	0.112	0.0014	
3768	0.8848	0.0014	
4347	0.0014	0.0014	●●●●●●●●
5046	0.0014	0.0014	●●●●●●●●
5364	0.0014	0.0014	●●●●●●●●

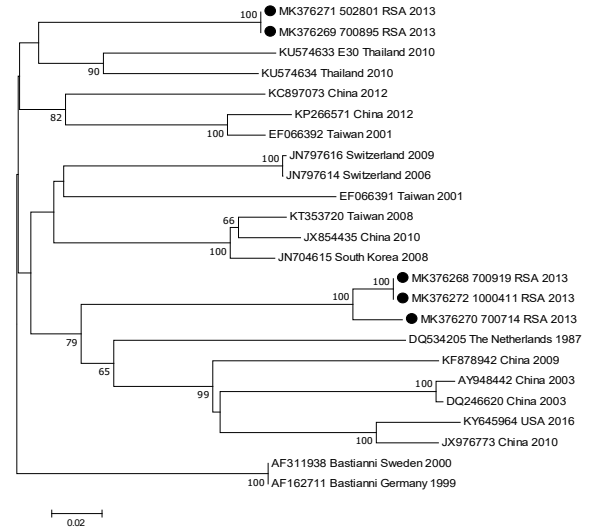
Figure 5.10: Topological incongruence report. LHS – left-hand side; RHS – right-hand side



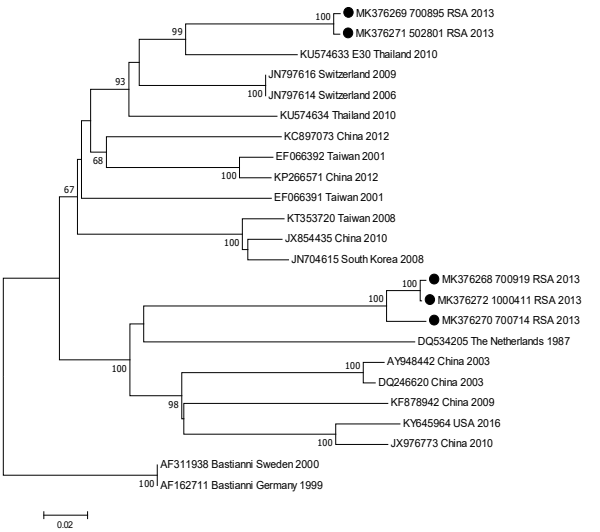
Nucleotide position 2886-4346



Nucleotide position 5046-5364



Nucleotide position 4347-5045



Nucleotide position 5364-6582

Figure 5.11: Phylogenetic relationships based on 4 partitions generated using the significant breakpoints across the echovirus 30 (E30) genome. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis.

5.4 Discussion

This is the first study to describe the use of SISPA-PCR for whole genome sequencing of enteroviruses, specifically E30 viruses. It is known that the use of SISPA might decrease the possibility of specific primer bias and provides the opportunity to detect emerging clades or viruses with mutations in the primer binding sites (57). In this study, enrichment for viral RNA (DNase treatment) decreased the double stranded cDNA concentration obtained during RACE-SISPA between 10-30 times. From these results, the sample with the highest yield did not produce the most reads. It suggests that if the input template concentrations are too low, it affects the efficiency of the NGS library preparation step. Addition of carrier RNA to the low concentration enriched RNA templates may have resulted in successful NGS, but we did not explore this. This finding is similar to that of Hall *et al*., (101) which highlights the need for more data to validate the effect of enrichment methods on different sample types, amplification, bioinformatics approaches and sequencing platforms.

The complete polyprotein nucleotide sequences of E30 strains from respiratory illness cases in South Africa, display a single biphyletic clade on genomic analysis with reference strains from mainly meningitis cases. The South African viruses clustered closely (94% bootstrap support) with strains from the Netherlands, China and USA - this clustering was sustained when analysis was performed on the sub-genomic P1region, which encodes the structural/capsid proteins of the E30 virus. These structural proteins are the primary antigens that induce protective immunity and are involved in receptor binding and therefore plays a role in immunity and infectivity (2). This indicates that the structural proteins, forming the outer capsid of the virus, were genetically distant from strains isolated from patients with aseptic meningitis.
Phylogenetic analysis was also performed on the P2 and P3 sub-genomic region – the non-structural proteins responsible for viral protease, shutoff of host protein synthesis, increased membrane permeability, inhibition of cellular secretory pathways, vesicle formation and nucleotide triphosphatase activity. For the non-structural sub-genomic P2 and P3 regions, South African strains did not cluster

together but form 2 distinct subclusters; two South African strains cluster closely with Asian strains from 2008 and 2010 of which two were isolated from respiratory cases from Thailand in 2010.

Complete full genome phylogenetic analysis supports the original designation of genotype k for the South African strains which was based on the partial VP1 region (Chapter 4); confirmed in this study with 100% bootstrap support during complete VP1 analysis. However further subgenomic analysis based on significant recombination breakpoints show distinct evolutionary histories for the five E30 strains which forms 2 distinct South African clades as also described in sub-genomic analysis of P2 and P3 regions.

Molecular epidemiology studies of E30 meningitis outbreaks have shown that point mutations give rise to substantial genetic diversity in the capsid genes, specifically the VP1 polypeptide (2). The unique clustering of the South African E30 strains from respiratory specimens could be due to specific mutations in the capsid genes, although this was not investigated as part of this study. Complete VP1 and full genome phylogenetic analysis support the designation of a new genotype for South African E30 strains and specifically being designated genotype k as proposed previously. The fact that there are so little whole genome E30 sequences available from respiratory samples, could also be due to the fact that not a lot of research has been done on E30 viruses as a cause of respiratory illness.

Similar clustering was observed in the genomic polyprotein and P1 sub-genomic regions, but clustering differed significantly between other sub-genomic regions, similar results have been reported in a number of other studies (102). This could be due to different evolutionary paths of specific genes, an example of a gene symbiosis system, where genes, not species, are the elementary units of evolution. Distinct grouping was confirmed that could be a result of the diversity of different subgenomic regions in E30 driven by different evolutionary paths. This study only looked at intra-genotypic recombination and not at inter-genotypic recombination

that are generally observed in the non-structural regions (97). It was previously shown that recombination increases with geographical separation and time with recombination half-live estimated at 3.1 years for E30 (2,103). It is recommended that more full genome E30 positive samples from South Africa be sequenced to complement the five full E30 genomes submitted. As more complete genome sequence data from circulating E30 become available, the role of recombination on the evolutionary pathways of E30 may become apparent.

From this study, five E30 whole genome sequences were submitted to the public domain, and this contribution will increase the pool of whole genome sequences from sub-Saharan Africa. The majority of E30 genome data in the public domain are from patients with aseptic meningitis identified in Asia, USA and Europe ranging from 1987 to 2012, which may explain the genetic distance of these viruses compared to viruses from respiratory samples in this study.

Ongoing molecular surveillance is important to help identify emerging virus strains of clinical significance and to better understand circulation patterns of these viruses.

Chapter 6

The Fraction of Rhinovirus Detections Attributable To Mild and Severe Respiratory Illness in a Setting of High Human Immunodeficiency Virus Prevalence, South Africa, 2013–2015

Reference:

Hellferscee O, Treurnicht FK, Walaza S, du Plessis M, von Gottberg A, Moyes J, Dawood H, Variava E, Pretorius M, Venter M, Cohen C, Tempia S. The fraction of rhinovirus detections attributable to mild and severe respiratory illness in a setting of high human immunodeficiency virus prevalence, South Africa, 2013–2015. Journal of Infectious Diseases 2019 May 5;219(11):1697-1704. Doi:10.1093/infdis/jiy725

6.1 Introduction

In the family *Picornaviridae*, the genus Enterovirus includes rhinovirus and enterovirus species (104). Rhinoviruses are amongst the most commonly detected viruses in patients with upper respiratory tract infection and they are considered to be associated, in most cases, with mild, self-limiting cold-like illnesses (41,42). However, in studies conducted in South Africa, rhinoviruses were detected in 25%-34% of upper respiratory tract specimens from patients hospitalised with SARI (21,105).

Recent advances in the development of sensitive PCR assays have considerably expanded the ability of laboratories to detect pathogens, especially viral agents. Nonetheless, establishing a clinical association between pathogen detection and illness remains challenging. Improving the understanding of the fraction of rhinoviruses detection associated with mild or severe respiratory illness would provide insights for the interpretation of surveillance data, including the burden of illness associated with rhinoviruses infection.

This study aimed to estimate the fraction of rhinovirus detection attributable to illness (hereafter defined as RV-AF) among HIV-infected and HIV-uninfected South African patients of different age groups separately during 2013-2015. This was done among outpatients with ILI and inpatients with severe respiratory illness (SRI). In addition, the proportion of ILI and SRI cases attributable to rhinoviruses infection after adjusting for the RV-AF were estimated.

6.2 Materials and Methods

6.2.1 Study design and population

Study samples were obtained from participants enrolled in the SRI (SARI and SCRI combined) and ILI surveillance programmes as well as asymptomatic individuals from June 2013 through July 2015 (described in section 2.1). For this study, 2012 samples were excluded as control recruitment started late and we therefore extended the study period to 2015 instead of 2014.

6.2.2 Sample selection and detection of rhinovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2-2.3.

Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay (14) (methods section 2.4.1-2.4.2) during 2013-2014. In 2015, a combination of the commercial Fast Track Diagnostics influenza virus and respiratory syncytial virus assay (Fast Track Diagnostics, Luxembourg) and Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) was used. Human bocavirus, human coronaviruses (OC43, HK1, NL63, 229E) and para-influenza virus 4 were only tested in 2015 and were excluded from the analysis. The platform was changed to a commercial assay because the in-house multiplex PCR was no-longer available and detection of all 10 pathogens were validated (with samples containing high and low viral loads) before implementing the commercial assay.

6.2.3 Statistical analysis

As described in section 2.5.

6.3 Results

6.3.1 Study population

Over the study period, 10,576 individuals were enrolled, of whom 10,134 (95.8%) had known age and available rhinovirus and HIV results, and were thus included for further analysis. From the 10,134 study participants, 2120 (20.9%) were controls, 3654 (36.1%) had ILI; and 4360 (43.0%) had SRI. Of the 4360 patients with SRI, 2735 (62.7%) had SARI and 1625 (37.3%) had SCRI. Children aged <5 years accounted for 38.0% (806/2120) of controls, 35.5% (1297/3654) of patients with ILI and 37.4% (1630/4360) of patients with SRI [55.6% (1520/2735) of patients with SARI; and 6.7% (110/1625) of patients with SCRI].

Overall, the HIV prevalence was 41.2% (874/2120) among controls (owing to enrolment criteria for controls); 25.0% (915/3654) among patients with ILI and 50.1% (2186/4360) among patients with SRI [37.1% (1015/2735) among patients with SARI; and 72.1% (1171/1625) among patients with SCRI ($p<0.001$)]. Across syndromes, the HIV prevalence was lowest among infants aged <1 year [1.5% (7/464) among those with ILI; 9.3% (96/1032) among those with SRI; 8.7% (84/961) among those with SARI; and 16.9% (12/71) among those with SCRI], and it was highest among patients aged 25–44 years [59.3% (581/980) among those with ILI, 90.5% (1221/1349) among those with SRI; 89.3% (509/570) among those with SARI; and 91.4% (712/779) among those with SCRI] (Table 6.1).

Table 6.1: HIV prevalence among outpatients with influenza-like illness and inpatients with severe (acute or chronic) respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013-December 2015

Age, year	Influenza-like illness	Severe respiratory illness	Severe acute respiratory illness	Severe chronic respiratory illness
	% (n/N)	% (n/N)	% (n/N)	% (n/N)
<1	1.5 (7/464)	9.3 (96/1032)	8.7 (84/961)	16.9 (12/71)
1–4	2.0 (17/833)	13.7 (82/598)	12.7 (71/559)	28.2 (11/39)
5–24	15.7 (152/970)	60.4 (198/328)	54.3 (102/188)	68.6 (96/140)
25–44	59.3 (581/980)	90.5 (1221/1349)	89.3 (509/570)	91.4 (712/779)
45–64	43.9 (154/351)	64.9 (535/824)	65.6 (225/343)	64.5 (310/481)
≥65	7.1 (4/56)	23.6 (54/229)	21.1 (24/114)	26.1 (30/115)
Total	25.0 (915/3654)	50.1 (2186/4360)	37.1 (1015/2735)	72.1 (1171/1625)

6.3.2 Rhinovirus detection and attributable fraction

In all age groups during the study period, rhinoviruses were detected in 23.2% (2350/10134) of specimens from enrolled study participants; 17.4% (368/2120) among controls (Table 6.2), 26.8% (979/3654) among patients with ILI (Table 6.3), and 23.0% (1003/4360) among patients with SRI ($p<0.001$) (Table 6.4) [27.5% (751/2735) among those with SARI (Table 6.5); and 15.5% (252/1625) among those with SCRI (Table 6.6)].

The overall RV-AF was 54.6% (95% CI: 47.5-60.8%) among patients with ILI (Table 6.3, Figure 6.1A) and 51.4% (95% CI: 43.4-58.3%) among patients with SRI (Table 6.4, Figure 6.2A) [54.3% (95% CI: 46.3-61.1%) among those with SARI (Table 6.5); and 37.5% (95% CI: 21.5-50.2%) among those with SCRI (Table 6.6)]. The RV-AF was significant among both HIV-infected and HIV-uninfected individuals with ILI (HIV-infected: 59.9%, 95% CI: 45.8-70.3% vs HIV-uninfected: 53.6%, 95% CI: 44.7-61.1%) or SRI (HIV-infected: 39.8%, 95% CI: 22.3-53.3% vs HIV-uninfected: 55.3%, 95% CI: 45.6-63.2%) (Tables 6.3 and 6.4, Figure 6.1B-C, Figure 6.2B-C). The RV-AF was also significant among HIV-infected and HIV-uninfected individuals aged <5 and ≥5 years with ILI or SRI, with the exception of HIV-infected children aged <5 years with SRI (Tables 6.3 and 6.4). Non-significant RV-AF were observed in some of the other age groups- and HIV-stratified analyses.

Overall, the RV AF-adjusted prevalence was 14.6% among patients with ILI (Table 6.3, Figure 6.1A) and 11.8% among patients with SRI (Table 6.4, Figure 6.2A) [14.9% among those with SARI (Table 6.5); and 5.8% among those with SCRI (Table 6.6)].

Table 6.2. Asymptomatic individuals positive for rhinovirus infection, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Age, year	All no. positive/no. total (%)	HIV-infected no. positive/no. total (%)	HIV-uninfected no. positive/no. total (%)
<1	98/339 (28.9%)	7/31 (22.6%)	91/308 (29.5%)
1–4	121/467 (25.9%)	53/205 (25.9%)	68/262 (26.0%)
5–24	94/592 (15.9%)	43/280 (15.4%)	51/312 (16.3%)
25–44	27/310 (8.7%)	17/189 (9.0%)	10/121 (8.3%)
45–64	18/280 (6.4%)	12/141 (8.5%)	6/139 (4.3%)
≥65	10/132 (7.6%)	3/28 (10.7%)	7/104 (6.7%)
<5	219/806 (27.2%)	60/236 (25.4%)	159/570 (27.9%)
≥5	149/1314 (11.3%)	75/638 (11.8%)	74/676 (10.9%)
All	368/2120 (17.4%)	135/874 (15.4%)	233/1246 (18.7%)

Table 6.3. Rhinoviruses detection and rhinoviruses attributable fraction among outpatients with influenza-like illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Outpatients with influenza-like illness									
	All			HIV-infected			HIV-uninfected		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%
<1	**48.1 (28.6 to 62.3)**	181/464 (39.0)	18.8	96.5 (0 to 99.9)	3/7 (42.9)	41.4	**46.7 (26.4 to 61.5)**	178/457 (38.9)	18.2
1–4	**40.8 (18.9 to 56.8)**	269/833 (32.3)	13.2	54.3 (0 to 86.0)	6/17 (35.3)	19.2	**40.1 (17.0 to 56.8)**	263/816 (32.2)	12.9
5–24	**53.9 (38.8 to 65.2)**	248/970 (25.6)	13.8	**40.9 (0.9 to 64.8)**	34/152 (22.4)	9.2	**58.8 (41.9 to 70.8)**	214/818 (26.2)	15.4
25–44	**68.7 (52.2 to 79.5)**	204/980 (20.8)	14.3	**66.0 (41.8 to 80.1)**	119/581 (20.5)	13.5	**72.6 (46.0 to 86.2)**	85/399 (21.3)	15.5
45–64	**73.5 (54.2 to 84.6)**	68/351 (19.4)	14.3	**70.5 (41.0 to 85.2)**	36/154 (23.4)	16.5	**76.7 (44.0 to 90.3)**	32/197 (16.2)	12.4
≥65	61.7 (0 to 85.6)	9/56 (16.1)	9.9	84.4 (0 to 98.8)	1/4 (25)	21.1	59.9 (0 to 86.0)	8/52 (15.4)	9.2
<5	**44.6 (30.7 to 55.7)**	450/1297 (34.7)	15.5	**68.6 (10.6 to 88.98)**	9/24 (37.5)	25.7	**43.6 (29.1 to 55.2)**	441/1273 (34.6)	15.1
≥5	**62.9 (54.4 to 69.8)**	529/2357 (22.4)	14.1	**59.5 (44.5 to 70.49)**	190/891 (21.3)	12.7	**64.4 (53.0 to 73.0)**	339/1466 (23.1)	14.9
All	**54.6 (47.5 to 60.8)**	979/3654 (26.8)	14.6	**59.9 (45.8 to 70.3)**	199/915 (21.7)	13.1	**53.6 (44.7 to 61.1)**	780/2739 (28.5)	15.3

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in non-stratified models.

Table 6.4. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe respiratory illness									
All				**HIV-infected**			**HIV-uninfected**		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%
<1	**38.8 (17.6 to 54.5)**	332/1032 (32.2)	12.5	18.6 (0 to 68.8)	27/96 (28.1)	5.2	**41.5 (20.0 to 57.3)**	305/936 (32.6)	13.5
1–4	**60.2 (45.8 to 70.7)**	241/598 (40.3)	24.3	20.3 (0 to 58.0)	30/82 (36.6)	7.4	**68.4 (54.6 to 77.9)**	211/516 (40.9)	28
5–24	**55.2 (37.2 to 68.0)**	96/328 (29.3)	16.2	**45.5 (13.5 to 65.6)**	53/198 (26.8)	12.2	**64.0 (41.3 to 78.0)**	43/130 (33.1)	21.2
25–44	**47.4 (18.8 to 65.9)**	219/1349 (16.2)	7.7	**51.1 (18.1 to 70.8)**	204/1221 (16.7)	8.5	30.4 (0 to 69.9)	15/128 (11.7)	3.6
45–64	**43.8 (5.0 to 66.7)**	92/824 (11.2)	4.9	26.5 (0 to 61.2)	62/535 (11.6)	3.1	**60.4 (5.0 to 83.5)**	30/289 (10.4)	6.3
≥65	30.5 (0 to 68.0)	23/229 (10.0)	3.1	-23.7 (0 to 71.7)	4/54 (7.4)	0	42.7 (0 to 76.7)	19/175 (10.9)	4.7
<5	**50.3 (38.6 to 59.9)**	573/1630 (35.2)	15.7	21.2 (0 to 53.8)	57/178 (32.0)	6.8	**56.0 (44.3 to 65.3)**	516/1452 (35.5)	19.9
≥5	**51.3 (38.7 to 61.3)**	430/2730 (15.8)	8.1	**44.8 (25.6 to 59.1)**	323/2008 (16.1)	7.2	**57.9 (39.9 to 70.5)**	107/722 (14.8)	8.6
All	**51.4 (43.4 to 58.3)**	1003/4360 (23.0)	11.8	**39.8 (22.3 to 53.3)**	380/2186 (17.4)	6.9	**55.3 (45.6 to 63.2)**	623/2174 (28.7)	1.9

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

Table 6.5. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe acute respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe acute respiratory illness									
	Overall			**HIV-infected**			**HIV-uninfected**		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%
<1	**41.1 (20.4 to 56.4)**	312/961 (32.5)	13.4	29.4 (0 to 73.3)	26/84 (31.0)	9.1	**43.2 (22.0 to 58.7)**	286/877 (32.6)	14.1
1–4	**60.5 (45.8 to 71.2)**	224/559 (40.1)	24.3	7.3 (0 to 53.1)	24/71 (33.8)	2.5	**69.2 (55.6 to 78.7)**	200/488 (41.0)	28.4
5–24	**61.8 (43.7 to 74.1)**	61/188 (32.4)	20.1	**43.6 (1.3 to 67.8)**	26/102 (25.5)	11.1	**73.7 (54.6 to 84.8)**	35/86 (40.7)	30
25–44	**55.8 (29.4 to 72.3)**	103/570 (18.1)	10.1	**59.6 (30.5 to 76.5)**	96/509 (18.9)	11.3	38.1 (0 to77.2)	7/61 (11.5)	3.9
45–64	**49.1 (8.5 to 71.7)**	40/343 (11.7)	5.7	14.3 (0 to 58.5)	24/225 (10.7)	1.5	**73.2 (30.3 to 89.7)**	16/118 (13.6)	9.9
≥65	27.7 (0 to 71.1)	11/114 (9.6)	2.6	0 (0 to 75.2)	0/24 (0)	0	50.0 (0 to 81.8)	11/90 (12.2)	6
<5	**51.6 (39.9 to 61.0)**	536/1520 (35.3)	18.2	16.4 (-45.3 to 51.9)	50/155 (32.3)	5.3	**57.2 (45.5 to 66.3)**	486/1365 (35.6)	20.4
≥5	**58.0 (45.7 to 67.6)**	215/1215 (17.7)	10.3	**45.4 (23.8 to 60.9)**	146/860 (17.0)	7.7	**68.8 (53.7 to 79.0)**	69/355 (19.4)	13.3
All	**54.3 (46.3 to 61.1)**	751/2735 (27.5)	14.9	**39.0 (19.3 to 54.0)**	196/1015 (19.3)	7.5	**59.5 (50.3 to 67.0)**	555/1720 (32.3)	19.2

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

Table 6.6. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe chronic respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe chronic respiratory illness									
	Overall			HIV-infected			HIV-uninfected		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%
<1	0 (0 to 45.0)	20/71 (28.2)	-0.5	0 (0 to 69.4)	1/12 (8.3)	-8.9	17.8 (0 to 57.9)	19/59 (32.2)	5.7
1–4	**54.9 (9.4 to 77.6)**	17/39 (43.6)	23.9	63.5 (0 to 90.1)	6/12 (50.0)	34.7	52.1 (0 to 79.3)	11/28 (39.3)	20.5
5–24	**39.6 (3.9 to 62.1)**	35/140 (25.0)	9.9	**46.6 (5.1 to 70.0)**	27/96 (28.1)	13.1	32.1 (0 to 70.6)	8/44 (18.2)	5.8
25–44	**41.6 (7.0 to 63.3)**	116/779 (14.9)	6.2	**43.7 (4.0 to 67.0)**	108/712 (15.2)	6.6	28.2 (0 to 73.3)	8/67 (11.9)	3.4
45–64	41.1 (0 to 66.6)	52/481 (10.8)	4.4	38.8 (0 to 69.4)	38/310 (12.3)	4.8	44.8 (0 to 78.8)	14/171 (8.2)	3.7
≥65	34.0 (0 to 72.5)	12/115 (10.4)	3.5	33.7 (0 to 85.2)	4/30 (13.3)	4.5	34.5 (0 to 77.1)	8/85 (9.4)	3.2
<5	26.8 (0 to 53.7)	37/110 (33.6)	9	26.4 (0 to 73.6)	7/24 (30.4)	8	34.4 (0 to 61.0)	30/87 (34.5)	11.9
≥5	**39.9 (21.5 to 54.0)**	215/1515 (14.2)	5.7	**44.6 (22.7 to 60.3)**	177/1148 (15.4)	6.9	33.7 (0 to 58.6)	38/367 (10.4)	3.5
All	**37.5 (21.5 to 50.2)**	252/1625 (15.5)	5.8	**42.9 (21.6 to 58.3)**	184/1172 (15.7)	6.7	**33.6 (6.34 to 53.0)**	68/454 (15.0)	5

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

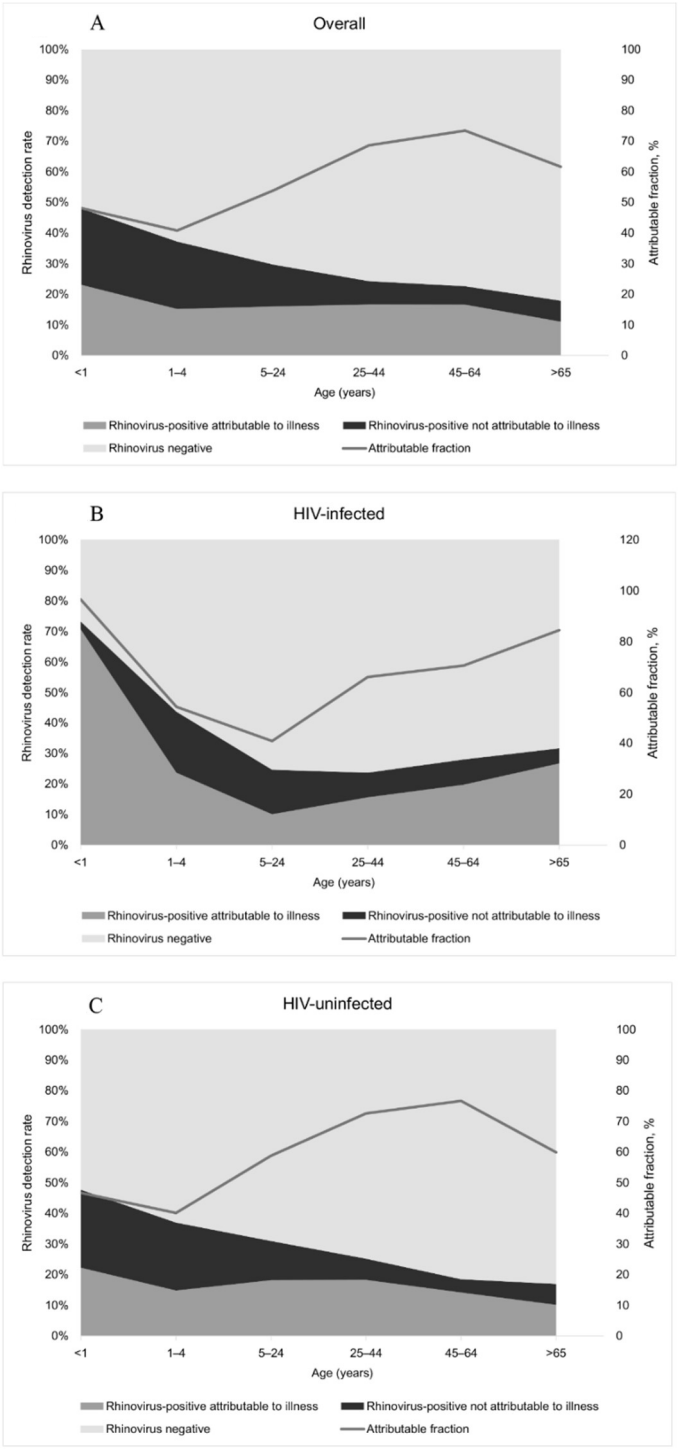


Figure 6.1: Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among outpatients with influenza-like illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected patients.

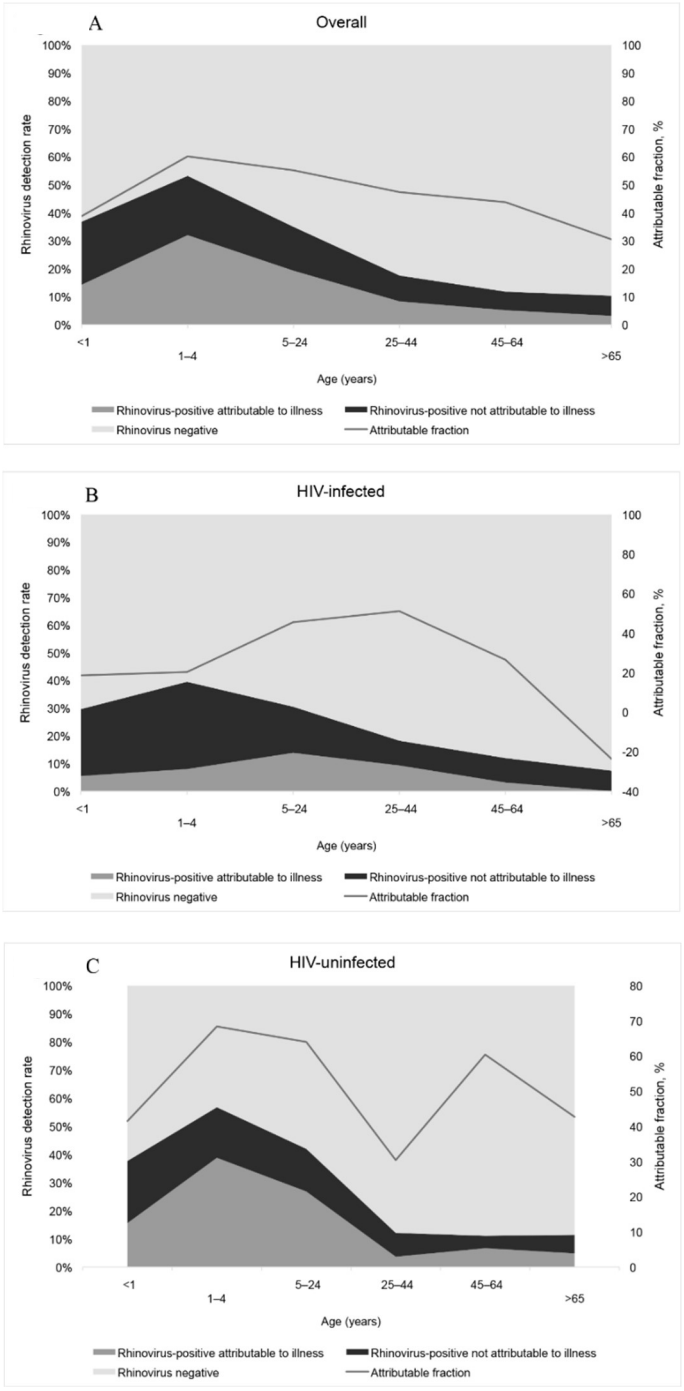


Figure 6.2: Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among patients hospitalised with severe respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected

6.4 Discussion

This study assessed the association between rhinovirus detection and mild or severe respiratory illness, among HIV-infected and HIV-uninfected patients of different age groups separately. Overall, rhinovirus was significantly associated with illness across syndromes and HIV status; however, this association was not significant in some age groups. The estimated detection rate attributable to illness reported in this study reflects a more accurate measure of the proportion of rhinovirus detections causing respiratory illness in both children and adults in South Africa, contrary to reporting viral detection rates alone. This study confirmed an overall significant RV-AF among patients with mild or severe respiratory illness as previously described in South Africa (21).

The overall RV-AF (ILI: 54.6%; SARI: 54.3%) found in this study were similar to those reported in a previous study conducted in South Africa (ILI: 52.0%; SARI: 46.9%), but was lower compared to those of other important respiratory pathogens such as influenza viruses (ILI: 93.3%; SARI: 86.3%) and respiratory syncytial virus (ILI: 63.1%; SARI: 83.7%) (21). Although this study found a significant RV-AF among HIV-infected and HIV-uninfected individuals across syndromes, only approximately half of the illness among RV-positive patients would be actually attributable to rhinovirus infection.

It was hypothesized that the elevated HIV infection prevalence in the South African setting may have been responsible for the significant attributable fraction found in a previous South African study (21); however, this study observed that rhinovirus detection was significantly associated with both mild and severe respiratory illness also among HIV-uninfected individuals. Studies from Kenya, Mali, Thailand and USA did not find an association of rhinovirus infection with illness (47–52,106); however, these studies had some limitations, including tests that combined rhinovirus and enterovirus detection and small samples sizes. The significant RV-AF found in this study among both HIV-infected and HIV-uninfected individuals may be due to increased power associated with the large sample size of this study compared to the above-mentioned studies. After

adjusting for the RV-AF the proportion of illness attributable to rhinovirus (ILI: 11.8%; SARI: 14.9%) was similar to or higher than those of influenza viruses (ILI: 14.2%; SARI: 4.9%) or respiratory syncytial virus (ILI: 5.9%; SARI: 16.7%) (21,107). This suggests that rhinovirus, whereas not always associated with illness when detected, may still be responsible for a substantial number of mild and severe respiratory illness cases.

This study has limitations that warrant discussion. Firstly, rhinovirus nucleic acids could be present during the pre- or post-syndromic phase of the infection and can be detected using PCR. In addition, controls after enrolment were not followed-up, and the development of undetected mild or severe illness for some individual cannot be excluded. The inclusion in the control group of individuals with rhinovirus -associated symptomatic illness during the pre- or post-syndromic phase would have resulted in an underestimation of the RV-AF. Secondly, the role of bacterial co-infections was not investigated in this analysis. Third, the small sample size in some of the age group- and HIV-stratified analyses may have hindered the ability to reach statistical significance in some of the groups. In addition, not having complete CD4+ count data, hindered the ability to assess the RV-AF among severely vs. mildly or not immunosuppressed HIV-infected individuals. Fourth, due to low numbers this study could not assess the RV-AF among patients with very severe illness, such as ICU admission or death. Last, no lower respiratory tract specimens were collected, hindering the ability to assess a differential rhinovirus positivity proportion between lower and upper respiratory tract specimens.

In conclusion, rhinoviruses were commonly identified among asymptomatic individuals but may still be associated with a substantial proportion of mild and severe respiratory illness across age groups and among HIV-infected and HIV-uninfected individuals. The proportion of illness attributable to rhinovirus after adjusting for the RV-AF was similar to or higher than those of other important pathogens such as influenza viruses and respiratory syncytial virus. This has

potential implications on the prioritization of pathogens for vaccines and/or therapeutics development

Chapter 7

Conclusions

Enteroviruses are known causes of respiratory illness, but the epidemiology and contribution has not been well described in South Africa. In this study, the epidemiology and genotypes associated with SARI in South African patients were determined. Enterovirus genotypes associated with SARI were compared to those from milder ILI and asymptomatic controls to determine if certain enterovirus genotypes were more prevalent in mild or severe patients with respiratory illness. In this study enteroviruses frequently contributed to SARI and ILI and that there was a substantial diversity in enterovirus genotypes circulating in South African patients hospitalised with SARI. EV-D68, an enterovirus genotype of public health importance, circulated at high frequencies among children less than 5 years of age but did not contribute to more severe disease and adverse outcome. Unfortunately, the surveillance system did not record symptoms of neurological disease or paralysis and therefore this study was unable to investigate a causative link between EV-D68 infection and neurological disease.

Furthermore, this study did not identify any association between disease severity and different enterovirus genotypes when comparing mild (ILI) and severe (SARI) cases to asymptomatic controls. However, some genotypes (E30, coxsackie virus B5, EV-D68) were more prevalent in symptomatic cases. This study also reported that E30 viruses which were the most prevalent in the respiratory illness cases formed a distinct genetic cluster upon phylogenetic analysis of the sub genomic VP1 region which I proposed to call genotype k. However, requirements for new virus clade designations require that genome sequencing is done.

Therefore, in this study the, SISPA-based NGS was employed to sequence the complete E30 genomes, predominantly from paediatric SARI cases. I successfully sequenced five E30 genomes which were submitted to GenBank; contributing to the pool of whole genome sequences available for E30, especially from sub-Saharan Africa. No evidence of recombination could be confirmed, although distinct grouping was observed that could be a result of the diversity of different subgenomic regions in E30.

Lastly we also wanted to determine the epidemiology of rhinovirus species from the genus Enterovirus and assess the proportion of rhinovirus positives that contribute to clinical illness in the context of a high HIV prevalence setting. Previous studies reported that that HIV infection caused increased severity and higher case fatality ratios, especially in children (6,7). We determined the attributable fraction of rhinovirus infections among HIV-infected and HIV-uninfected patients with mild and severe respiratory illness in South Africa during 2013 to 2015. Rhinoviruses are the most prevalent species from the genus Enterovirus identified among patients hospitalised with SARI, but they are also like enterovirus commonly identified in healthy individuals. For this reason, it was decided to determine the fraction of rhinovirus detection associated with mild or severe respiratory illness. An overall significant attributable fraction of rhinoviruses among patients with mild or severe respiratory illness was confirmed, which suggest that rhinoviruses, whereas not always associated with illness, may still be responsible for a substantial number of mild and severe respiratory illness cases. The proportion of illness attributable to rhinoviruses (ILI: 11.8%; SARI: 14.9%) were similar to or higher than other pathogens of public health importance such as influenza viruses (ILI: 14.2%; SARI: 4.9%) and respiratory syncytial virus (ILI: 5.9%; SARI: 16.7%) (21).

In order to determine if other factors such as viral load or host characteristics play a role in enterovirus-associated disease, further studies will need to be conducted to increase the knowledge regarding the epidemiology of enteroviruses.

Currently, there is no vaccine or antiviral therapy available for non-polio enteroviruses and by understanding the contribution of enterovirus genotypes to severe illness this will allow for informed decision making on potential candidate vaccines and development of therapeutic interventions.

To conclude, study findings show that (i) enterovirus genotypes detected in mild versus severe disease do not differ, (ii) that outbreak-associated EV strains like EV-D68 are circulating in South Africa and therefore ongoing surveillance is important, (iii) that E30 was the most prevalent EV strains and that E30 strains from respiratory illness cases in South Africa belongs to a newly proposed genotype k and lastly (iv) we conclude that rhinovirus species although found in asymptomatic illness cases may still contribute to a substantial number symptomatic illness cases across all age groups regardless of HIV status.

REFERENCES

1. Viral Meningitis [Internet]. Centres for Disease Control and Prevention. 2018. Available from: https://www.cdc.gov/meningitis/viral.html

2. Bailly JL, Mirand A, Henquell C, Archimbaud C, Chambon M, Charbonné F, et al. Phylogeography of circulating populations of human echovirus 30 over 50 years: Nucleotide polymorphism and signature of purifying selection in the VP1 capsid protein gene. Infect Genet Evol. 2009;9(4):699–708.

3. Pons-Salort M, Parker EPK, Grassly NC. The epidemiology of non-polio enteroviruses. Curr Opin Infect Dis [Internet]. 2015;28(5):479–87.

4. Dasaraju P V, Liu C. Chapter 93 Infections of the Respiratory System. In: Medical Microbiology. 1996.

5. GBD 2015 LRI Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect Dis. 2017 Nov;17(11):1133–61.

6. Gray DM, Zar HJ. Community-acquired pneumonia in HIV-infected children: a global perspective. Curr Opin Pulm Med. 2010 May;16(3):208–16.

7. Cohen C, Moyes J, Tempia S, Groom M, Walaza S, Pretorius M, et al. Severe influenza-associated respiratory infection in high HIV prevalence setting, South Africa, 2009-2011. Emerg Infect Dis. 2013;19(11):1766–74.

8. Tiveljung-Lindell A, Rotzén-Östlund M, Gupta S, Ullstrand R, Grillner L, Zweygberg-Wirgart B, et al. Development and implementation of a molecular diagnostic platform for daily rapid detection of 15 respiratory viruses. J Med Virol. 2009;81(1):167–75.

9. Brittain-Long R, Nord S, Olofsson S, Westin J, Anderson LM, Lindh M. Multiplex real-time PCR for detection of respiratory tract infections. J Clin Virol. 2008;41(1):53–6.

10. Pretorius MA, van Niekerk S, Tempia S, Moyes J, Cohen C, Madhi SA, et al. Replacement and positive evolution of subtype A and B respiratory syncytial virus G-protein genotypes from 1997-2012 in South Africa. J Infect Dis. 2013 Dec;208 Suppl:S227-37.

11. Cohen AL, Hellferscee O, Pretorius M, Treurnicht F, Walaza S, Madhi S, et al. Epidemiology of influenza virus types and subtypes in South Africa, 2009–20121. Emerg Infect Dis [Internet]. 2014;20(7):1149–56.

12. Haynes LM. Progress and challenges in RSV prophylaxis and vaccine development. J Infect Dis. 2013 Dec;208 Suppl:S177-83.

13. WHO. Pneumonia Fact sheet N°331 [Internet]. web page WHO. 2016. Available from: http://www.who.int/mediacentre/factsheets/fs331/en/

14. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ, et al. ICTV Virus Taxonomy Profile: Picornaviridae. J Gen Virol. 2017 Oct;98(10):2421–2.

15. Fields BN, Knipe DM, Howley PM. Fields Virology, 5th Edition [Internet]. Vol. 2, Fields Virology. 2007. 3177 p.

16. Alexander He, Koch G, Mountain Im, Sprunt K, Van Damme O. Infectivity of ribonucleic acid of poliovirus on HeLa cell mono-layers. Virology. 1958 Feb;5(1):172–3.

17. Hober D, Sauter P. Pathogenesis of type 1 diabetes mellitus: Interplay between enterovirus and host. Nature Reviews Endocrinology. 2010 May;6(5):279-89.

18. Oberste MS, Maher K, Kilpatrick DR, Flemister MR, Brown BA, Pallansch MA. Typing of human enteroviruses by partial sequencing of VP1. J Clin Microbiol. 1999;37(5):1288–93.

19. Palmenberg AC, Spiro D, Kuzmickas R, Wang S, Djikeng A, Rathe JA, et al. Sequencing and analyses of all known human rhinovirus genomes reveal structure and evolution. Science 2009 Apr 3;324(5923):55-9.

20. McIntyre CL, Knowles NJ, Simmonds P. Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. J Gen Virol. 2013 Aug;94(Pt 8):1791-806

21. Pretorius MA, Tempia S, Walaza S, Cohen AL, Moyes J, Variava E, et al. The role of influenza, RSV and other common respiratory viruses in severe acute respiratory infections and influenza-like illness in a population with a high HIV sero-prevalence, South Africa 2012-2015. J Clin Virol. 2016;75:21–6.

22. Carrillo-Tripp M, Shepherd CM, Borelli IA, Venkataraman S, Lander G, Natarajan P, et al. VIPERdb2: an enhanced and web API enabled relational database for structural virology. Nucleic Acids Res. 2009 Jan;37(Database issue):D436-42.

23. Jartti T, Lehtinen P, Vuorinen T, Koskenvuo M, Ruuskanen O. Persistence of Rhinovirus and Enterovirus RNA after Acute Respiratory Illness in Children. J Med Virol. 2004 Apr;72(4):695-9

24. Smura T, Ylipaasto P, Klemola P, Kaijalainen S, Kyllönen L, Sordi V, et al. Cellular tropism of human enterovirus D species serotypes EV-94, EV-70, and EV-68 in vitro: Implications for pathogenesis. J Med Virol. 2010;82(11):1940–9.

25. Wen H, Si L, Yuan X, Hao S, Gao F, Chu F, et al. Complete genome sequencing and analysis of six enterovirus 71 strains with different clinical phenotypes. Virol J [Internet]. 2013;10:115.

26. Pretorius MA, Madhi SA, Cohen C, Naidoo D, Groome M, Moyes J, et al. Respiratory viral coinfections identified by a 10-Plex real-time reverse-transcription polymerase chain reaction assay in patients hospitalized with severe acute respiratory illness-South Africa, 2009-2010. J Infect Dis. 2012

Dec 15;206 Suppl 1:S159-65..

27. Van Leer-Buter CC, Poelman R, Borger R, Niesters HGM. Newly identified Enterovirus C genotypes, identified in the Netherlands through routine sequencing of all Enteroviruses detected in clinical materials from 2008 to 2015.. J Clin Microbiol. 2016 Sep;54(9):2306-14

28. Tokarz R, Hirschberg DL, Sameroff S, Haq S, Luna G, Bennett AJ, et al. Genomic analysis of two novel human enterovirus C genotypes found in respiratory samples from Peru. J Gen Virol. 2013 Jan;94(Pt 1):120-7.

29. Renois F, Bouin A, Andreoletti L. Enterovirus 68 in pediatric patients hospitalized for acute airway diseases. J Clin Microbiol. 2013;51(2):640–3.

30. Xiang Z, Xie Z, Wang Z, Ren L, Xiao Y, Li L, et al. Human enterovirus genotype C104, China. Emerg Infect Dis. 2013 Apr;19(4):689-91

31. Xiang Z, Tsatsral S, Liu C, Li L, Ren L, Xiao Y, et al. Respiratory infection with enterovirus genotype C117, China and Mongolia. Emerg Infect Dis. 2014 Jun;20(6):1075-7

32. Puenpa J, Mauleekoonphairoj J, Linsuwanon P, Suwannakarn K, Chieochansin T, Korkong S, et al. Prevalence and characterization of enterovirus infections among pediatric patients with hand foot mouth disease, herpangina and influenza like illness in Thailand, 2012. PLoS One. 2014 Jun 2;9(6):e98888

33. Owens B. Rare enterovirus continues to circulate in North America. Vol. 384, Lancet (London, England). England; 2014. p. 1250.

34. Tokarz R, Firth C, Madhi SA, Howie SRC, Wu W, Sall AA, et al. Worldwide emergence of multiple clades of enterovirus 68. J Gen Virol. 2012;93(PART 9):1952–8.

35. Böttcher S, Prifert C, Weißbrich B, Adams O, Aldabbagh S, Eis-Hübinger AM, et al. Detection of enterovirus D68 in patients hospitalised in three tertiary university hospitals in Germany, 2013 to 2014. Eurosurveillance.

2016;21(19).

36. Gong YN, Yang SL, Shih SR, Huang YC, Chang PY, Huang CG, et al. Molecular evolution and the global reemergence of enterovirus D68 by genome-wide analysis. Med (United States). 2016 Aug;95(31):e4416

37. Imamura T, Fuji N, Suzuki A, Tamaki R, Saito M, Aniceto R, et al. Enterovirus 68 among children with severe acute respiratory infection, the Philippines. Emerg Infect Dis. 2011 Aug;17(8):1430-5

38. Henquell C, Mirand A, Richter J, Schuffenecker I, Böttiger B, Diedrich S, et al. Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission. J Virol [Internet]. 2013;87(22):12249–59.

39. Khetsuriani N, Lamonte-Fowlkes A, Oberst S, Pallansch MA. Enterovirus surveillance--United States, 1970-2005. MMWR Surveill Summ Morb Mortal Wkly report Surveill Summ / CDC. 2006;55(8):1–20.

40. Kyriakopoulou Z, Pliaka V, Amoutzias GD, Markoulatos P. Recombination among human non-polio enteroviruses: implications for epidemiology and evolution. Virus Genes. 2015 Apr;50(2):177–88.

41. Mertz D, Alawfi A, Pernica JM, Rutherford C, Luinstra K, Smieja M. Clinical severity of pediatric respiratory illness with enterovirus D68 compared with rhinovirus or other enterovirus genotypes. CMAJ. 2015 Nov;187(17):1279–84.

42. Harvala H, McIntyre CL, McLeish NJ, Kondracka J, Palmer J, Molyneaux P, et al. High detection frequency and viral loads of human rhinovirus species A to C in fecal samples; diagnostic and clinical implications. J Med Virol. 2012 Mar;84(3):536–42.

43. Royston L, Tapparel C. Rhinoviruses and Respiratory Enteroviruses: Not as Simple as ABC. Viruses. 2016 Jan;8(1).

44. Tapparel C, Junier T, Gerlach D, Van Belle S, Turin L, Cordey S, et al. New respiratory enterovirus and recombinant rhinoviruses among circulating picornaviruses. Emerg Infect Dis. 2009 May;15(5):719-26

45. Miller EK, Lu X, Erdman DD, Poehling KA, Zhu Y, Griffin MR, et al. Rhinovirus-associated hospitalizations in young children. J Infect Dis. 2007 Mar 15;195(6):773-81

46. Arden KE, McErlean P, Nissen MD, Sloots TP, Mackay IM. Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. J Med Virol. 2006 Sep;78(9):1232-40.

47. Fry AM, Lu X, Olsen SJ, Chittaganpitch M, Sawatwong P, Chantra S, et al. Human rhinovirus infections in rural Thailand: epidemiological evidence for rhinovirus as both pathogen and bystander. PLoS One. 2011 Mar;6(3):e17780.

48. Hammitt LL, Kazungu S, Morpeth SC, Gibson DG, Mvera B, Brent AJ, et al. A preliminary study of pneumonia etiology among hospitalized children in Kenya. Clin Infect Dis [Internet]. 2012 Apr;54 Suppl 2:S190-9.

49. Benet T, Sylla M, Messaoudi M, Sanchez Picot V, Telles J-N, Diakite A-A, et al. Etiology and factors associated with pneumonia in children under 5 years of age in mali: A Prospective case-control study. PLoS One. 2015;10(12):e0145447.

50. Breiman RF, Cosmas L, Njenga M, Williamson J, Mott JA, Katz MA, et al. Severe acute respiratory infection in children in a densely populated urban slum in Kenya, 2007-2011. BMC Infect Dis. 2015 Feb;15:95.

51. Self WH, Williams DJ, Zhu Y, Ampofo K, Pavia AT, Chappell JD, et al. Respiratory viral detection in children and adults: comparing asymptomatic controls and patients with community-acquired pneumonia. J Infect Dis. 2016 Feb;213(4):584–91.

52. Singleton RJ, Bulkow LR, Miernyk K, DeByle C, Pruitt L, Hummel KB, et al. Viral respiratory infections in hospitalized and community control children in Alaska. J Med Virol. 2010;82(7):1282–90.

53. Wolter N, Cohen C, Tempia S, Madhi SA, Venter M, Moyes J, et al. HIV and influenza virus infections are associated with increased blood pneumococcal load: a prospective, hospital-based observational study in South Africa, 2009-2011. J Infect Dis. 2014 Jan;209(1):56–65.

54. Tempia S, Walaza S, Viboud C, Cohen AL, Madhi SA, Venter M, et al. Mortality associated with seasonal and pandemic influenza and respiratory syncytial virus among children <5 years of age in a high HIV prevalence setting--South Africa, 1998-2009. Clin Infect Dis. 2014 May;58(9):1241–9.

55. Tempia S, Walaza S, Viboud C, Cohen AL, Madhi SA, Venter M, et al. Deaths associated with respiratory syncytial and influenza viruses among persons >/=5 years of age in HIV-prevalent area, South Africa, 1998-2009(1). Emerg Infect Dis. 2015 Apr;21(4):600–8.

56. Rames E, Macdonald J. Evaluation of MinION nanopore sequencing for rapid enterovirus genotyping. Virus Res. 2018 May;252:8–12.

57. Goya S, Valinotto LE, Tittarelli E, Rojo GL, Nabaes Jodar MS, Greninger AL, et al. An optimized methodology for whole genome sequencing of RNA respiratory viruses from nasopharyngeal aspirates. PLoS One. 2018 Jun 25;13(6):e0199714

58. Greninger AL. A decade of RNA virus metagenomics is (not) enough. Virus Res. 2018 Jan 15;244:218-229

59. Oberste MS, Maher K, Pallansch MA. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. J Virol. 2004 Jan;78(2):855-67.

60. Santti J, Hyypiä T, Kinnunen L, Salminen M. Evidence of recombination among enteroviruses. J Virol. 1999 Oct;73(10):8741-9.

61. Tan L Van, Tuyen NTK, Thanh TT, Ngan TT, Van HMT, Sabanathan S, et al. A generic assay for whole-genome amplification and deep sequencing of enterovirus A71. J Virol Methods. 2015 Apr;215–216:30–6.

62. Hirst GK. Genetic recombination with Newcastle disease virus, polioviruses, and influenza. Cold Spring Harb Symp Quant Biol. 1962;27:303-9.

63. Messacar K, Schreiner TL, Maloney JA, Wallace A, Ludke J, Oberste MS, et al. A cluster of acute flaccid paralysis and cranial nerve dysfunction temporally associated with an outbreak of enterovirus D68 in children in Colorado, USA. Lancet (London, England). 2015 Apr;385(9978):1662–71.

64. Allan Nix W, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. J Clin Microbiol. 2006;44(8):2698–704.

65. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

66. Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: More models, new heuristics and parallel computing. Nature Methods. 2012 Jul 30;9(8):772

67. Guindon S, Gascuel O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst Biol. 2003 Oct;52(5):696-704.

68. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24(8):1596–9.

69. Hall BG. Building phylogenetic trees from molecular data with MEGA. Mol Biol Evol. 2013;30(5):1229–35.

70. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, et al. Viral genome sequencing by random priming methods. BMC Genomics. 2008 Jan;9:5-2164-9–5.

71. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012 Jul 24;13:341.

72. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870-4.

73. Madden TL, Tatusov RL, Zhang J. Applications of network BLAST server. Methods in enzymology. 1996;266:131-41.

74. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015 May 26;1(1):vev003.

75. Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol. 2006 Oct;23(10):1891-901.

76. Delport W, Poon AF, Frost SD, Pond SLK. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. 2010 Oct;26(19):2455–7.

77. Zhou H, Alexander DH, Sehl ME, Sinsheimer JS, Sobel EM, Lange K. Penalized regression for genome-wide association screening of sequence data. Pac Symp Biocomput [Internet]. 2011;106–17.

78. Meijer A, Van Der Sanden S, Snijders BEP, Jaramillo-Gutierrez G, Bont L, Van Der Ent CK, et al. Emergence and epidemic occurrence of enterovirus 68 respiratory infections in The Netherlands in 2010. Virology [Internet]. 2012;423(1):49–57.

79. Renois F, Bouin A, Andreoletti L. Enterovirus 68 in pediatric patients hospitalized for acute airway diseases. J Clin Microbiol. 2013;51(2):640–3.

80. Meijer A, Benschop KS, Donker GA, van der Avoort HG. Continued seasonal circulation of enterovirus D68 in the Netherlands, 2011-2014. Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull. 2014;19(42).

81. Reiche J, Böttcher S, Diedrich S, Buchholz U, Buda S, Haas W, et al. Low-level circulation of enterovirus D68 associated acute respiratory infections, Germany, 2014. Emerg Infect Dis. 2015;21(5):837–41.

82. Tokarz R, Firth C, Madhi SA, Howie SRC, Wu W, Sall AA, et al. Worldwide emergence of multiple clades of enterovirus 68. J Gen Virol. 2012;93(PART 9):1952–8.

83. Du J, Zheng B, Zheng W, Li P, Kang J, Hou J, et al. Analysis of enterovirus 68 strains from the 2014 North American outbreak reveals a new clade, indicating viral evolution. PLoS One. 2015 Dec 2;10(12):e0144208.

84. Todd AK, Hall RJ, Wang J, Peacey M, McTavish S, Rand CJ, et al. Detection and whole genome sequence analysis of an enterovirus 68 cluster. Virol J. 2013 Apr 2;10:103.

85. Schuster JE, Miller JO, Selvarangan R, Weddle G, Thompson MT, Hassan F, et al. Severe enterovirus 68 respiratory illness in children requiring intensive care management. J Clin Virol. 2015 Sep;70:77-82.

86. Tapparel C, Siegrist F, Petty TJ, Kaiser L. Picornavirus and enterovirus diversity with associated human diseases. Infect Genet Evol. 2013 Mar;14:282-93.

87. Hellferscee O, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, et al. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza

Other Respi Viruses. 2017 May;11(3):211–9.

88. Yarmolskaya MS, Shumilina EY, Ivanova OE, Drexler JF, Lukashev AN. Molecular epidemiology of echoviruses 11 and 30 in Russia: Different properties of genotypes within an enterovirus serotype. Infect Genet Evol. 2015;30:244–8.

89. Liu N, Jia L, Yin J, Wu Z, Wang Z, Li P, Hao R et al. An outbreak of aseptic meningitis caused by a distinct lineage of coxsackievirus B5 in China. Int j Infect Dis [Internet]. 2014;23:101–4.

90. Papa A, Dumaidi K, Franzidou F, Antoniadis A. Genetic variation of coxsackie virus B5 strains associated with aseptic meningitis in Greece. Clin Microbiol Infect. 2006 Jul;12(7):688–91.

91. Midgley CM, Jackson MA, Selvarangan R, Turabelidze G, Obringer E, Johnson D, et al. Severe respiratory illness associated with enterovirus D68 - Missouri and Illinois, 2014. MMWR Morb Mortal Wkly Rep. 2014 Sep 12;63(36):798-9.

92. Patel MC, Wang W, Pletneva LM, Rajagopala S V, Tan Y, Hartert T V, et al. Enterovirus D-68 infection, prophylaxis, and vaccination in a novel permissive animal model, the Cotton Rat (Sigmodon hispidus). PLoS One. 2016;11(11):e0166336.

93. Janes VA, Minnaar R, Koen G, van Eijk H, Dijkman-de Haan K, Pajkrt D, et al. Presence of human non-polio enterovirus and parechovirus genotypes in an Amsterdam hospital in 2007 to 2011 compared to national and international published surveillance data: A comprehensive review. Vol. 19, Eurosurveillance. 2014. p. 1–9.

94. Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, et al. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. J Med Virol. 2017 Oct;89(10):1759–67.

95. Dyda A, Stelzer-Braid S, Adam D, Chughtai AA, Macintyre CR. The association between acute flaccid myelitis (AFM) and enterovirus D68 (EV-D68) – what is the evidence for causation? Eurosurveillance. 2018 Jan;23(3).

96. National Center for Biotechnology Information [Internet]. GenBank. 2018 [cited 2018 Jun 11]. Available from: https://www.ncbi.nlm.nih.gov/

97. Kyriakopoulou Z, Amoutzias GD, Dimitriou TG, Tsakogiannis D, Mossialos D, Markoulatos P. Intra- and inter-serotypic recombinations in the 5′ UTR-VP4 region of Echovirus 30 strains. Arch Virol. 2018 Feb;163(2):365-375.

98. Dimitrov KM, Sharma P, Volkening JD, Goraichuk I V., Wajid A, Rehmani SF, et al. A robust and cost-effective approach to sequence and analyze complete genomes of small RNA viruses. Virol J. 2017 Apr 7;14(1):72.

99. Chrzastek K, Lee D hun, Smith D, Sharma P, Suarez DL, Pantin-Jackwood M, et al. Use of Sequence-Independent, Single-Primer-Amplification (SISPA) for rapid detection, identification, and characterization of avian RNA viruses. Virology. 2017 Sep;509:159-166.

100. Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, et al. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. J Med Virol. 2017;89(10):1759–67.

101. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J Virol Methods. 2014 Jan;195:194-204.

102. Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE, Ilonen J. Recombination in circulating enteroviruses. J Virol. 2003 Oct;77(19):10423-31.

103. McWilliam Leitch EC, Cabrerizo M, Cardosa J, Harvala H, Ivanova OE, Kroes ACM, et al. Evolutionary dynamics and temporal/geographical correlates of recombination in the human enterovirus echovirus types 9, 11, and 30. J Virol [Internet]. 2010/07/07. 2010 Sep;84(18):9292–300.

104. Pallansch MA, Roos R. Enteroviruses: Polioviruses, coxsackieviruses, echoviruses and new enteroviruses. In: Knipe DM, Howley MP, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE. editiors. Fields Virology. 5th ed. Philadelphia: Lippincott Williams & Wilkens; 2007. p. 839-93

105. Pretorius MA, Tempia S, Treurnicht FK, Walaza S, Cohen AL, Moyes J, et al. Genetic diversity and molecular epidemiology of human rhinoviruses in South Africa. Influenza Other Respi Viruses. 2014 Sep;8(5):567–73.

106. Onyango CO, Welch SR, Munywoki PK, Agoti CN, Bett A, Ngama M, et al. Molecular epidemiology of human rhinovirus infections in Kilifi, coastal Kenya. J Med Virol. 2012 May;84(5):823–31.

107. Tempia S, Walaza S, Moyes J, Cohen AL, von Mollendorf C, McMorrow ML, et al. Attributable fraction of influenza virus detection to mild and severe respiratory illnesses in HIV-infected and HIV-uninfected patients, South Africa, 2012-2016. Emerg Infect Dis. 2017 Jul;23(7):1124–32.

Epidemiology and genotypes of Enteroviruses in patients with respiratory illness in South Africa

Orienka Hellferscee

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirement for the degree of Doctor of Philosophy

Johannesburg, 2019

Declaration

I, Orienka Hellferscee, declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

6 December 2019, Johannesburg

International and national presentations from this study

1. **Orienka Hellferscee**, Florette Treurnicht, Stefano Tempia, Ebrahim Variava, Halima Dawood, Kathy Kahn, Shabir Madhi, Adam Cohen, Cheryl Cohen and Marietjie Venter. Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients Hospitalised with Severe Acute Respiratory Illness (Oral presentation)
    - 1st International Respiratory Pathogens Conference, Singapore 1-5 September 2015

2. **Orienka Hellferscee**, Stefano Tempia, Ebrahim Variava, Halima Dawood, Shabir A. Madhi, Mignon du Plessis, Florette K. Treurnicht, Cheryl Cohen. Enterovirus Genotypes Detected Among Patients With Severe Acute Respiratory Illness, Influenza-like Illness And Asymptomatic Individuals In South Africa, 2012-2014 (Poster presentation)
    - 5th African Network for Influenza Surveillance and Epidemiology (ANISE) meeting, Kigali Rwanda, 8 - 11 March 2016
    - WITS research day 2016, Johannesburg, 1 September 2016

3. **Orienka Hellferscee,** Florette Treurnicht, Sibongile Walaza, Mignon du Plessis, Anne von Gottberg, Jocelyn Moyes, Halima Dawood, Ebrahim Variava, Marthi Pretorius, Marietjie Venter, Cheryl Cohen, Stefano Tempia. The attributable fraction of rhinovirus detection to mild and severe respiratory illness among HIV-infected and HIV-uninfected South African patients, 2013-2015" (Poster presentation)
    - 6th African Network for Influenza Surveillance and Epidemiology (ANISE) meeting, Antananarivo Madagascar, 19-21 March 2018
    - WITS research day 2018, Johannesburg, 6 September 2018

Publications directly contributing to this thesis and role of the student in these studies

Data from all participants are from respiratory illness surveillance conducted over a six-year period (2009-2015) for which I was responsible for performing and overseeing virology laboratory testing of surveillance samples from 2011 for the duration of the study.

In addition, I conducted all characterisation testing, phylogenetic analyses and statistical analyses for the publications included in this thesis and drafted and revised the published manuscripts.

1. **Hellferscee O**, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, Cohen AL, Pretorius M, Cohen C, Madhi SA, Venter M. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza and Other Respiratory Viruses. 2017 May; 11(3):211-219

2. **Hellferscee O**, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, Madhi SA, du Plessis M, Cohen C, Treurnicht FK. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. Journal of Medical Virology 2017 Oct; 89(10):1759-1767

3. **Hellferscee O**, Treurnicht FK, Walaza S, du Plessis M, von Gottberg A, Moyes J, Dawood H, Variava E, Pretorius M, Venter M, Cohen C, Tempia S. The fraction of rhinovirus detections attributable to mild and severe respiratory illness in a setting of high human immunodeficiency virus prevalence, South Africa, 2013–2015. Journal of Infectious Diseases 2019 May 5;219(11):1697-1704

4. **Hellferscee O**, Janse van Vuren P, du Plessis M, Cohen C, Treurnicht FK. Complete genome sequence analysis of human echovirus 30 isolated from patients with acute respiratory illness in South Africa, 2013. (In preparation)

Ethics clearance certificate

UNIVERSITY OF THE
WITWATERSRAND
JOHANNESBURG

R14/49 Ms O Hellferscee

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180386

NAME: (Principal Investigator)	Ms O Hellferscee
DEPARTMENT:	National Institute for Communicable Diseases Centre for Respiratory Diseases and Meningitis National Health Laboratory Service
PROJECT TITLE:	Epidemiology and genotypes of enteroviruses in patients with respiratory illness in South Africa
DATE CONSIDERED:	06/04/2018
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M140824 and antecedents
SUPERVISOR:	Drs C Cohen, M du Plessis and F Treurnicht
APPROVED BY:	Professor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	06/04/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

Abstract

Globally, lower respiratory tract infections (LRTI) are the leading cause of death in children aged less than 5 years. This study focuses on enteroviruses, which are a known cause of respiratory illness but whose epidemiology and molecular characterization have not been well described in South Africa.

The first aim of this study was to describe enterovirus genotypes circulating among South African patients hospitalized with severe respiratory illness (SRI). Upper respiratory samples were tested for the presence of enteroviruses using polymerase chain reaction (PCR) and characterized using a partial VP1 gene segment through Sanger sequencing. Enteroviruses were identified in 8.2% of patients with SRI. We demonstrated a high diversity, 17 genotypes, from all four enterovirus species circulating in South Africa.

The second aim was to use Sanger sequencing to characterize the enterovirus genotypes circulating among South African patients with influenza-like illness (ILI), SRI, and asymptomatic individuals. We found high genetic diversity (33 genotypes) but no association of enterovirus species with disease severity; however, some genotypes (echovirus 30 (E30), coxsackie virus B5, and EV-D68) were more prevalent in symptomatic cases. The E30 strains formed a new cluster proposed to be named genotype k.

The third aim described a novel approach for next generation sequencing of E30, based on sequence-independent single primer amplification (SISPA). Complete full genome phylogenetic analysis supports the designation of genotype k, and distinct grouping was confirmed that could be a result of the diversity of different subgenomic regions in E30 driven by different evolutionary paths. Five E30 whole genomes were submitted to GenBank, contributing to global data.

The fourth aim was to determine the fraction of rhinovirus (the most commonly identified species of Enterovirus) detection associated with mild or severe

respiratory illness. Rhinovirus was detected in 17.4% (368/2120), 26.8% (979/3654) and 23.0% (1003/4360) of controls, ILI and SRI, respectively. The overall RV-AF (ILI: 54.6%; severe acute respiratory illness (SARI): 54.3%) were similar to those reported in a previous study conducted in South Africa (ILI: 52.0%; SARI: 46.9%), but was lower compared to those of other important respiratory pathogens such as influenza viruses (ILI: 93.3%; SARI: 86.3%) and respiratory syncytial virus (ILI: 63.1%; SARI: 83.7%) (21). Even though the attributable fraction was lower than influenza and RSV, the attributable burden of rhinoviruses (RV-AF) was equivalent to influenza and RSV, because of the high rhinovirus detection rate.

In conclusion, we found (i) no difference in enterovirus genotypes detected in mild versus severe disease; (ii) that outbreak-associated EV strains like EV-D68 are circulating in South Africa; (iii) that E30 strains from respiratory illness cases in South Africa belong to a newly proposed genotype k; and (iv) that rhinovirus species contribute to mild and severe respiratory illness.

Acknowledgements

Alle eer aan my hemelse Pappa vir krag en verstand - “Ek is tot alles instaat deur U wat my krag gee”

My gesin – Alttden/Kayden/Laylah – dankie vir al die tyd wat julle my gegun het en vir al die koffie aandra. Girls, my “boek” is uiteindelik klaar en Lief, ek belowe, hierdie is die laaste ene ☺

My ouers – Pappie/Moemie, julle eindelose oproepe, ondersteuning en aanmoediging is van onskatbare waarde

My swes – Debs – dankie dat jy my nommer een cheerleader is!

Darinka/Jo-Andrie/Cardia – Buddies, dankie vir al die ondersteuning en aanmoediging, julle het ‘n spesiale plekkie in my hart

Amelia – my werk mamma – dankie vir al jou ondersteuning en dat jy soms meer werk moes aanvat sodat ek kon opskryf

My supervisors – I could not have asked for a better team:

- Cheryl/Florette/Mignon – thank you for believing in me as a scientist. Thank you for all your kind words and motivation and always making time to brainstorm, I am truly blessed to be mentored by the best
- Anne – thank you for all your guidance and teaching me how to think outside the box

Stefano/Meredith – thank you for all the reviews as part of writing group and Stefano thank you for teaching me statistics

Table of contents

Declaration……..... i
International and national presentations from this study……...……. ii
Publications directly contributing to this thesis and role of the student in these studies……..…….. iii
Ethics clearance certificate…….. v
Abstract…….. vi
Acknowledgements…….. viii
Table of contents……...……..... ix
List of figures……... xiv
List of tables…….. xx
List of abbreviations……... xxii

Chapter 1 Literature Review……... 1

1.1 Global overview of respiratory tract infection……..... 1
1.2 Respiratory tract infections in South Africa……... 2
1.3 Family *Picornaviridae*…… 3
1.3.1 Genus Enterovirus…… 5
1.3.2 Enterovirus species……... 6
1.3.2.1 Clinical manifestations and seasonality…… 6
1.3.2.2 Enterovirus D68…… 7
1.3.2.3 Echovirus 30…… 8
1.4 Rhinovirus species……..... 9
1.4.1 Clinical manifestations and seasonality…… 9
1.4.2 Does rhinovirus contribute to the burden of severe respiratory illness?..... 10
1.5 Viral genomics, viral evolution and recombination events ……... 11
1.6 Aims and objectives……. 13

Chapter 2 Materials and Methods…… 15

2.1 Description of the surveillance programmes and case definitions…… 15
2.1.1 Severe acute respiratory illness (SARI) surveillance…….. 15
2.1.2 Severe chronic respiratory illness (SCRI) surveillance…….. 16
2.1.3 Severe respiratory illness (SRI) surveillance…….. 16
2.1.4 Influenza-like illness (ILI) surveillance…….. 17
2.1.5 Asymptomatic individuals…….. 17
2.2 Study procedures……... 17
2.3 Description of sample collection……... 18
2.4 Laboratory methods and procedures……. 19
2.4.1 Total nucleic acid extraction……... 19
2.4.2 Laboratory methods for enterovirus and rhinovirus……...…… 20
2.4.3 Sanger sequencing of enterovirus VP1 gene region……... 21
2.4.3.1 Assembly and phylogenetic analysis of enterovirus partial VP1... 23
2.4.4 Enterovirus isolation……...… 23
2.4.5 Enterovirus genome sequencing……..… 23
2.4.5.1 Viral enrichment and sequence independent single primer amplification (RACE-SISPA) of echovirus 30……...… 23
2.4.5.1.1 Viral enrichment……. 23
2.4.5.1.2 Rapid amplification of cDNA (RACE)……...… 24
2.4.5.1.3 Sequence independent single primer amplification (SISPA)……...… 24
2.4.5.2 Next generation sequencing on the MiSeq platform…….. 26
2.4.5.3 Data analysis: Trimming, mapping of the reads and assembly….. 26
2.4.6 Phylogenetic analysis of echovirus 30 whole genome…… 27
2.5 Statistical data analysis……. 27

Chapter 3 Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients with Severe Acute Respiratory Illness, 2009-2011…… 30

3.1 Introduction……..... 30
3.2 Materials and Methods……..….... 31
3.2.1 Study design and population……...… 31

3.2.2 Sample selection and detection of enterovirus infection……. 31
3.2.3 Amplification and sequencing of enteroviruses partial VP1…..…..… 32
3.2.4 Statistical analysis……... 32
3.3 Results……..... 32
3.3.1 Clinical and epidemiological characteristics of patients hospitalised with enterovirus-associated SARI……...… 32
3.3.2 Molecular epidemiology of enterovirus species A to D…….. 33
3.3.3 Enterovirus serotypes and circulation of EV-D68 clades during 2009-2011……...… 36
3.4 Discussion……..… 41

Chapter 4 Enterovirus Genotypes among Patients with Severe Acute Respiratory Illness, Influenza-Like Illness and Asymptomatic Individuals in South Africa, 2012-2014……..… 43

4.1 Introduction……...… 43
4.2 Materials and Methods…….. 44
4.2.1 Study design and population……...… 44
4.2.2 Sample selection and detection of enterovirus infection……..… 44
4.2.3 Amplification and sequencing of enteroviruses…….. 44
4.2.4 Statistical analysis……...… 45
4.3 Results……... 45
4.3.1 Epidemiology of the study population……..... 45
4.3.2 Phylogenetic comparison of enterovirus genotypes identified during June 2012 – July 2014……. 46
4.3.3 Enterovirus group and genotypes…… 49
4.3.4 Molecular evolution of echovirus 30, coxsackie virus B5 and EV-D68 genotypes…….. 53
4.3.4.1 Echovirus 30……...… 53
4.3.4.2 Coxsackie virus B5……..… 53
4.3.4.3 EV-D68……...… 56
4.4 Discussion……..… 58

Chapter 5 Genome Analysis of Human Echovirus 30 Isolated from Patients with Acute Respiratory Illness in South Africa, 2013……..… 61

5.1 Introduction……... 61
5.2 Material and Methods..…… 63
5.2.1 Study design and population……..…….…... 63
5.2.2 Sample selection and detection of enterovirus infection…….….. 63
5.2.3 Amplification and sequencing of enteroviruses…… 63
5.2.4 Enterovirus isolation and genome sequencing…….. 63
5.3 Results…….. 64
5.3.1 Enterovirus isolation……..… 64
5.3.2 Characteristics of the echovirus 30 positive cases……. 64
5.3.3 Genome sequencing……... 65
5.3.4 Complete genome analysis of echovirus 30 samples…… 66
5.3.5 Phylogenetic analysis…… 68
5.3.5.1 Phylogenetic analysis of complete genomes…...……... 68
5.3.5.2 Phylogenetic analysis of sub-genomic P1 region……... 70
5.3.5.3 Phylogenetic analysis of sub-genomic P2 and P3 regions …… 71
5.3.5.4 VP1 sub-genomic analysis……...……... 73
5.3.6 Investigation of E30 genomes for recombination……..… 74
5.4 Discussion……..… 76

Chapter 6 The Fraction of Rhinovirus Detections Attributable To Mild and Severe Respiratory Illness in a Setting of High Human Immunodeficiency Virus Prevalence, South Africa, 2013–2015……... 79

6.1 Introduction……...… 79
6.2 Materials and Methods…….…. 80
6.2.1 Study design and population……..… 80
6.2.2 Sample selection and detection of rhinovirus infection…….. 80
6.2.3 Statistical analysis……..… 81
6.3 Results…….. 81
6.3.1 Study population…… 81

6.3.2 Rhinoviruses detection and attributable fraction……... 83
6.4 Discussion……...….. 91

Chapter 7 Conclusions…….…… 94

References……..... 97

List of figures

Figure	Title	Page
Figure 1.1	Negative-contrast electron micrograph of enterovirus particles; the bar represents 100 nm (courtesy of Ann C. Palmenberg)……	1
Figure 1.2	Diagrammatic presentation of the enterovirus capsid and genome structure (17)…….....	4
Figure 1.3	Pictures of the different human enterovirus species (family *Picornaviridae*, genus enterovirus). Species, type, virus name abbreviation and the virion's average diameter (Æ) are indicated (22)……..……..	6
Figure 2.1	Map of provinces in South Africa – indicating the two surveillance sites (courtesy: Sibongile Walaza, Centre for Respiratory Diseases and Meningitis, NICD)…….....	15
Figure 3.1	Molecular characterization of enterovirus clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implemented in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Black dots = South African samples 2009-2011……....	34
Figure 3.2	Epidemiologic graphs showing: (A) Distribution of enterovirus in South Africa during 2009-2011 and (B) Distribution of enterovirus species detected in South Africa during 2009-2011……	35
Figure 3.3	Molecular characterization of EV-D68 clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed	

using neighbor-joining methods as implicated in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Scale bar indicates number of nucleotide substitutions per site. Black dots = South African strains 2009-2011; Open dots = South African strains 2000-2001. Clades B1, B3, C A1 and partial B2 and A2 have been compressed to show comparison of South African strains to strains from other countries (country and year in blocks). Total number of sequences = 312 (representative strains from 2000-2016)……..... 40

Figure 4.1 Monthly number and overall detection rate of enterovirus (EV)-positive cases among controls (A), influenza-like illness (B) and severe acute respiratory patients (C), Klerksdorp and Pietermaritzburg, South Africa, 2012-2014…… 48

Figure 4.2 Phylogenetic analysis of enterovirus genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denotes serotypes identified in SARI patients, those with grey closed circles denotes genotypes identified in ILI patients while those with open circles denoted genotypes identifies in asymptomatic individuals. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree……... 50

Figure 4.3 The genotypes and species (EV-A to EV-D) of enteroviruses identified in patients with severe acute respiratory illness (SARI), influenza-like illness (ILI) and asymptomatic individuals (C) from June 2012 to July 2014, South Africa…….. 52

Figure 4.4 Distribution of enterovirus (EV) species (A-D) in South Africa during June 2012 to July 2014; dnt=did not genotype……..... 52

Figure 4.5 Phylogenetic analysis of echovirus 30 genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site……..... 54

Figure 4.6 Phylogenetic analysis of coxsackie virus B5 genogroups by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site…… 55

Figure 4.7 Phylogenetic analysis of EV-D68 lineages by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Sequences with open circles are EV-D68 strains previously identified from South Africa. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site……... 57

Figure 5.1 Genome structure of echovirus 30 (permission from ViralZone) (VPg – Viral protein genome linked; IRES – internal ribosomal entry site; RdRp – RNA dependent RNA polymerase)…… 62

Figure 5.2 Electrophoresis run from the Bioanalyzer indicating smear between 300-1000bp……... 65

Figure 5.3 Schematic representation of genome coverage after reference mapping against the E30 prototype, Bastianni strain (AF311938)……... 67

Figure 5.4 Entropy plot indicating the difference of the South African echovirus 30 strains across the polyprotein……. 66

Figure 5.5 Genomic (6580 bp) relationship of 5 South African echovirus 30 (E30) strains compared to 19 other E30 strains available in GenBank database. Genome sequences were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstraps replicated). The scale bar represents the genetic distance. Bootstrap values >65% are shown. ● South African strains from this study……. 69

Figure 5.6 Phylogenetic relationships based on the complete P1 sub genomic region (2580 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study…….. 70

Figure 5.7 Phylogenetic relationships based on the complete P2 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study……. 71

Figure 5.8 Phylogenetic relationships based on the complete P3 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study…….. 72

Figure 5.9 Phylogenetic relationships based on the complete VP1 genomic region (876 bp) of echovirus 30 (E30) strains. Five South African and 58 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis. Bold line indicates the proposed genotype k……..... 73

Figure 5.10 Topological incongruence report. LHS – left-hand side; RHS – right-hand side……. 74

Figure 5.11 Phylogenetic relationships based on 4 partitions generated using the significant breakpoints across the echovirus 30 (E30) genome. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis……..…….. 75

Figure 6.1 Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among outpatients with influenza-like illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected patients……. 88

Figure 6.2 Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among patients hospitalised with severe respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected……..... 89

List of tables

Table	Title	Page
Table 3.1	Clinical and epidemiological characteristics of patients hospitalised with enterovirus associated severe acute respiratory illness at surveillance sites, South Africa, 2009-2011……..	37
Table 3.2	Demographic and clinical characteristics of ten patients in whom EV-D68 was identified in South Africa…….....	39
Table 4.1	Association of enterovirus species (EV-A to EV-D) with influenza-like illness (ILI) and severe acute respiratory illness (SARI) compared with asymptomatic controls, South Africa, 2009-2010 (bold indicate factors significant at p<0.05)……	51
Table 5.1	Concentration of double stranded cDNA generated using RACE-SISPA comparing enriched and non-enriched viral samples as well as total reads generated with MiSeq……..	64
Table 6.1	HIV prevalence among outpatients with influenza-like illness and inpatients with severe (acute or chronic) respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013-December 2015……..	81
Table 6.2	Asymptomatic individuals positive for rhinovirus infection, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……...	83
Table 6.3	Rhinoviruses detection and rhinoviruses attributable fraction among outpatients with influenza-like illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……	84

Table 6.4 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015…….. 85

Table 6.5 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe acute respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015……... 86

Table 6.6 Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe chronic respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015…… 87

List of abbreviations

%	percentage
<	less than
≥	greater than or equal to
°C	degree Celsius
µg/ml	microgram per microliter
µl	microliter
Æ	average diameter
AF	attributable fraction
Ag/Ab	antigen and antibody
BLAST	Basic Local Alignment Search Tool
bp	base pairs
CD4+	T-cell CD4 T-helper cells
cDNA	complementary DNA
CI	confidence interval
Ct	cycle threshold
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E30	Echovirus 30
e.g	for example
EIA	enzyme immuno assay
ELISA	enzyme-linked immunosorbent assay
EV-A	enterovirus species A
EV-B	enterovirus species B
EV-C	enterovirus species C
EV-D	enterovirus species D
EV-D68	enterovirus D68
g	relative centrifugal force
GARD	Genetic Algorithm for Recombination Detection
HEp-2	Human epithelial type 2
HIV	human immunodeficiency virus
HPD	high probability density
i.e	in other words
ICU	intensive care unit
ILI	influenza-like illness
IRES	Type I internal ribosomal entry site
LRTI	lower respiratory tract infection
mM	millimolar
mRNA	messenger ribonucleic acid

NGS	next generation sequencing
NHLS	National Health Laboratory Service
NICD	National Institute for Communicable Diseases
nm	Nanometer
nt	nucleotides
NP/OP	combined nasopharyngeal and oropharyngeal specimen
OR	odds ratio
ORF	open reading frame
PCR	polymerase chain reaction
pmol	picomole
qRT-PCR	real-time reverse transcription polymerase chain reaction
RACE-SISPA	rapid amplification of cDNA end sequence independent single primer amplification
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
RV-A	rhinovirus species A
RV-AF	rhinovirus attributable fraction
RV-B	rhinovirus species B
RV-C	rhinovirus species C
SARI	severe acute respiratory illness
SCRI	severe chronic respiratory illness
SDS	sodium dodecyl sulfate
SISPA	sequence independent single primer amplification
SRI	severe respiratory illness
TNA	total nucleic acid
U	unit
UNICEF	United Nations Children's Fund
USA	United States of America
UTR	untranslated region
V	volts
VPg	genomic Viral Protein
vs.	versus
WHO	World Health Organization
µM	micromolar

Chapter 1

Literature Review

1.1 Global overview of respiratory tract infection

Lower respiratory tract infections (LRTI) are inflammation of the airways or pulmonary tissue, resulting from an infection below the level of the larynx and include conditions like bronchiolitis, bronchitis, laryngotracheobronchitis (croup) and pneumonia (4). LRTIs are a substantial public health problem and a leading cause of illness and death in people of all ages. LRTI's cause more than 2.75 million deaths worldwide, making this the 5th leading cause of death overall and the leading infectious cause of death in children less than 5 years of age (5). Despite substantial reductions in LRTI mortality in many countries, the burden remains high, particularly in areas of low sociodemographic development and has increased in some populations, particularly the extreme of ages – very young and elderly people (5).

Estimates of the global burden of LRTI's will be improved by more high-quality data on mortality, morbidity and aetiologies, especially in sub-Saharan Africa where the burden is highest and data are scarce. Improvements in diagnostics for LRTI's aetiologies, including those appropriate for children younger than five years of age to better understand the contribution of each aetiology to the LRTI burden, will help guide targeted interventions such as vaccination (5).

1.2 Respiratory tract infections in South Africa

LRTI is the most common cause of hospitalisation in human immunodeficiency virus (HIV)-infected children in sub-Saharan Africa and HIV infection is associated with increased severity of LRTI and higher case fatality ratios (6). The relative contribution of viral and bacterial aetiologies to the syndrome of pneumonia also varies by HIV status. Data from a severe acute respiratory illness (SARI) surveillance system, a prospective hospital based sentinel surveillance system, estimates the overall incidence of LRTI in South African children less than five years of age to be 2530-3173 per 100,000 with the highest incidence in children aged less than one year of age (annual range 8446-10532 per 100,000). The risk of hospitalisation with LRTI was up to three times higher in HIV-infected compared to HIV-uninfected children (7).

Influenza-like illness (ILI) in outpatients causes large numbers of illness episodes in South Africa and is a major reason for health facility usage in low and middle-income countries, contributing to the public health burden in terms of absenteeism from work, health-care and economic cost. Respiratory viruses associated with acute LRTI and ILI include influenza virus A and B, respiratory syncytial virus, parainfluenza virus types 1, 2 and 3, adenovirus, human metapneumovirus, enterovirus and rhinovirus (8,9).

Two of these viruses, influenza virus and respiratory syncytial virus, have been well characterized as causes of pneumonia in South Africa (10,11) and data from these studies have been used to guide vaccine policy and candidate vaccine strain selection (12). This fits well with the World Health Organization (WHO) and United Nations Children's Fund (UNICEF) integrated global action plan for pneumonia which aims to accelerate pneumonia control with a combination of interventions to protect, prevent and treat pneumonia in children (13). However, a number of other viruses has been suggested to cause pneumonia or ILI and have candidate vaccines in development, but these are not well described in South Africa. This includes viruses in the family *Picornaviridae*, genus Enterovirus – enteroviruses and rhinoviruses.

1.3 Family *Picornaviridae*

The family *Picornaviridae* consists of small, icosahedral non-enveloped viruses, 30-32 nm (14) (Figure 1.1). Picorna is derived from the prefix “pico” which designates a very small unit of measurement (10^{-12}) and RNA – therefore directly translated meaning very small RNA. The genome consists of a single-stranded non-segmented positive sense RNA of approximately 7500 nucleotides, containing of a single open reading frame (ORF). The ORF encoding a polyprotein flanked on either side by untranslated regions (UTRs) which contains regions of secondary structure, which are essential to genome function (14).

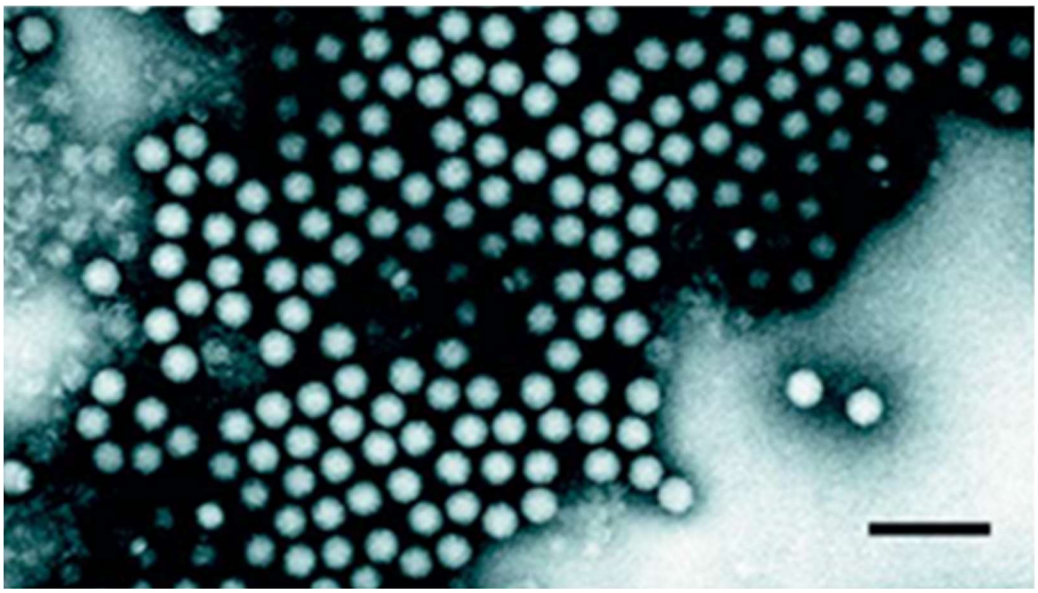


Figure 1.1: Negative-contrast electron micrograph of enterovirus particles; the bar represents 100 nm (courtesy of Ann C. Palmenberg).

The polyprotein is initially processed into three precursor proteins: P1, P2 and P3. The P1 genomic region encodes four capsid or structural proteins (VP1, VP2, VP3 and VP4) while the P2 and P3 genomic regions encode seven nonstructural proteins (2A, 2B, 2C, 3A, 3B, 3C and 3D) associated with replication (2,14,15). The capsid is made up of 60 copies of each of the four capsid proteins assembled into the icosahedral virion structure, and only the VP1, VP2 and VP3 proteins are exposed on the capsid surface (15) (Figure 1.2).

The viral genomic RNA has a viral protein (VPg) at its 5` end instead of a -methylated nucleotide cap structure. The long 5' UTR contains a type I internal ribosomal entry site (IRES), which allows direct translation of the polyprotein. The ORF also contains a poly(A) tail located after the 3' terminal sequence (2,14). The virion RNA is infectious, serving as both the genome and the viral mRNA (16).

The orthologous (genes in different species that evolved from a common ancestral gene by speciation) proteins VP2, VP3, VP1, 2C, 3C and 3D are conserved for the *Picornaviridae* family and may be used for sequence alignments between viruses, whereas VP4, 2A, 2B, 3A and 3B are highly divergent among the members of different picornavirus genera. None of the viral proteins are glycosylated (14). The family comprises 35 genera and human enteroviruses and rhinoviruses are members of the genus Enterovirus in the family *Picornaviridae* (14).

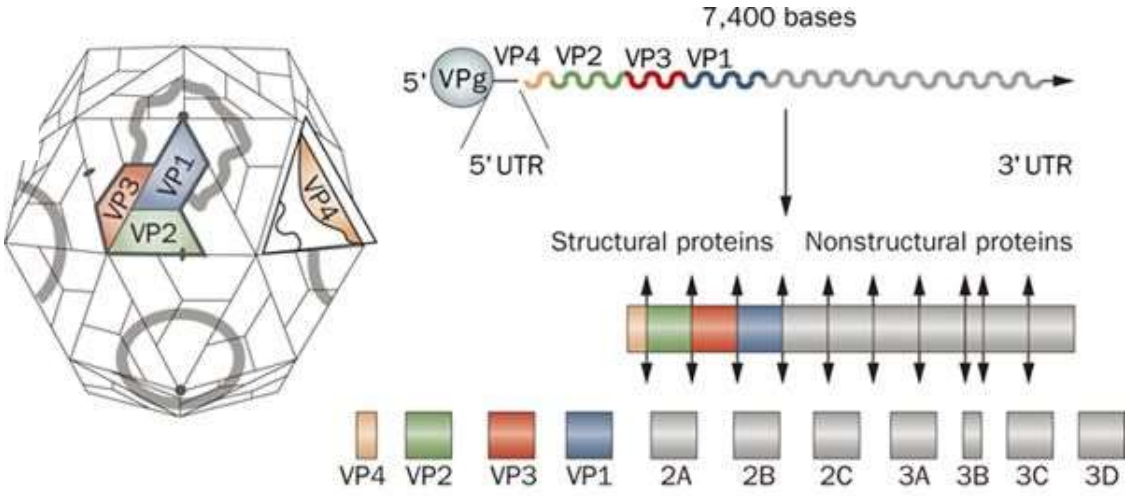


Figure 1.2: Diagrammatic presentation of the enterovirus capsid and genome structure (17).

1.3.1 Genus Enterovirus

Entero comes from the Greek word “enteron” which means intestine, which is an appropriate name since enteroviruses multiply primarily in the gastrointestinal tract or the upper respiratory tract or sometimes both, but they can also multiply in other tissues, e.g. nerve and muscle. Infections with enteroviruses are generally cytolytic, but persistent infections are common (14). Clinical manifestations include the common cold, mild meningitis, encephalitis, myelitis, myocarditis and conjunctivitis although some infections may frequently be asymptomatic (14). Transmission is horizontal, mainly faecal-oral, fomite or airborne routes (14).

Because of the high sequence diversity within the VP1 capsid region and genotype specific neutralization profiles, more than 300 virus types have been characterized in the genus Enterovirus (2). The genus genetically consists of about 75 enterovirus genotypes and 100 rhinovirus genotypes that cluster into 7 species (4 enterovirus species namely enterovirus A [EV-A], enterovirus B [EV-B], enterovirus C [EV-C], enterovirus D [EV-D] and 3 rhinovirus species namely rhinovirus A [RV-A], rhinovirus B [RV-B] and rhinovirus C [RV-C]) that infect humans (14) (Figure 1.3). Other types and genotypes include enterovirus E: 5 types, enterovirus F: 7 types, enterovirus G: 20 types, enterovirus H: 1 type, enterovirus I: 1 type, enterovirus J: 6 types, and at least 3 types that await classification. Four enterovirus species (A to D) are human pathogens but other enterovirus species (G to L) have been reported from animals (i.e. swine, monkey and rodent) (14,18–20).

The 5′-UTR of human rhinoviruses is shorter than that of enteroviruses, due to a deletion of approximately 100 nucleotides (nt) between the IRES and the translation start site (14). Human enteroviruses and rhinoviruses share identical genomic organization, have similar functional RNA secondary structures and are classified in the same genus because of their high sequence homology (14). However, despite their common genomic features, these two groups of viruses have different phenotypic characteristics including acid stability: virions of most

enterovirus species are stable at pH 3.0, while those of the rhinovirus species are unstable below pH 5-6 (14).

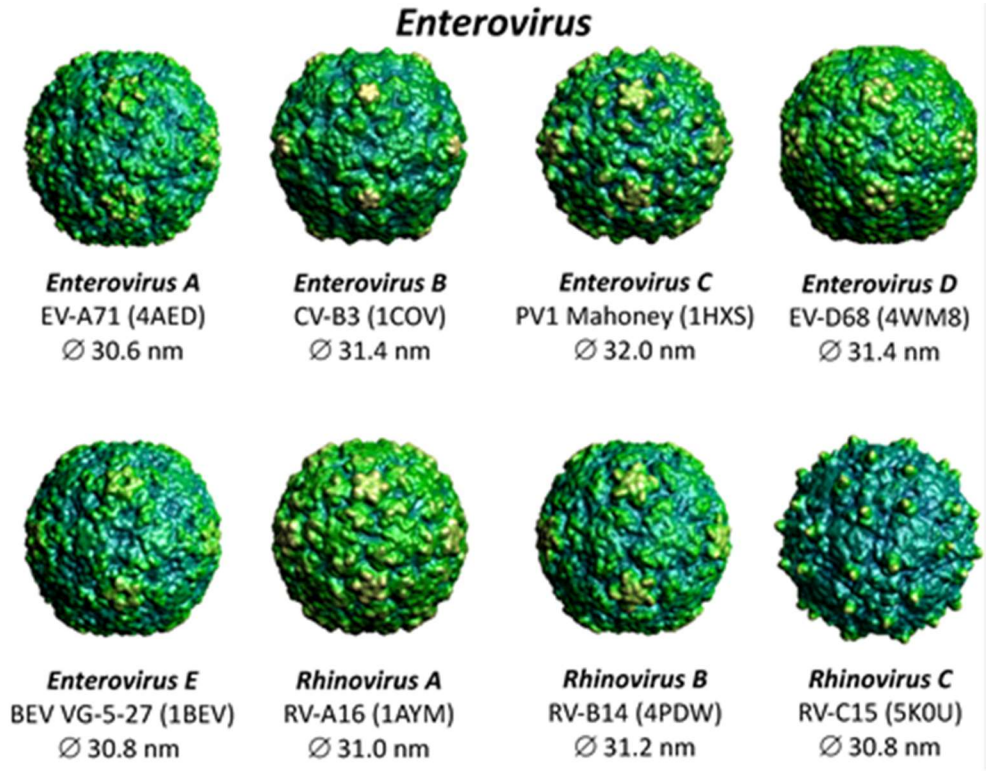


Figure 1.3: Pictures of the different human enterovirus species (family *Picornaviridae*, genus enterovirus). Species, type, virus name abbreviation and the virion's average diameter (Æ) are indicated (22).

1.3.2 Enterovirus species

1.3.2.1 Clinical manifestations and seasonality

Most human enteroviruses enter the body via the alimentary or respiratory tract, and spread by both faecal-oral and respiratory routes, although a few exceptions to the usual modes of transmission exist. Transmission of viral particles between humans occurs mainly via direct contact or fomites. Enterovirus RNA disappears over 2 to 5 weeks from the nasal cavity after an acute respiratory infection, and 97% of enterovirus-positive patients test negative for RNA by reverse transcription polymerase chain reaction (RT-PCR) 2 to 3 weeks after onset of illness (23). Enterovirus positivity, however, remained detectable longer than

culture positivity, suggesting that RNA may be remnants of past infections, without active replication.

Although the majority of enterovirus infections are subclinical, acute and clinical illnesses can present as including upper and lower respiratory illness, aseptic meningitis, encephalitis, acute haemorrhagic conjunctivitis, acute flaccid paralysis, myocarditis, and neonatal sepsis-like disease (24). Respiratory symptoms are mostly similar to a mild cold, yet in some cases can cause SARI requiring hospitalisation and may even be fatal (25). Enteroviruses are detected throughout the year, but have some peak activity in summer and autumn (26). Molecular techniques have demonstrated that new enterovirus genotypes causing respiratory illness have emerged or re-emerged (such as enterovirus-D68 [EV-D68], enterovirus 104, enterovirus 109, enterovirus 117 and coxsackie virus A21) (27–31). A study from Thailand in 2012 detected 11 different enterovirus genotypes from three species including EV-A (coxsackie virus A5, A6, A8, A10 and EV71), EV-B (coxsackie virus A9, B1, B2, echovirus 7 and echovirus 16) and EV-C (coxsackie virus A21) in paediatric patients with ILI (32).

No studies at the time of writing this thesis had determined whether different genotypes of enteroviruses are associated with mild (ILI) or severe acute respiratory illness (SARI) in other countries, and especially in South Africa, a country with high HIV prevalence.

1.3.2.2 Enterovirus D68

Recent reports of re-emerging lineages causing severe respiratory illness, some with fatal outcomes, were associated with EV-D68 infection in the Unites States of America (USA) (33). EV-D68 belongs to species EV-D and, until recently, sporadic outbreaks have been reported in Japan, Philippines, The Netherlands and the USA (34).

Reports of severe, fatal disease associated with EV-D68 infection in the USA renewed interest in enterovirus as the cause of SARI (33). EV-D68 infects primary target tissues of the respiratory tract directly, and occasionally infects the central nervous system (24). Bayesian phylogenetic analysis of EV-D68 strains globally, revealed the presence of three primary clades (A, B and C) although other literature have separated EV-D68 into clusters (1, 2 and 3) and lineages (1 and 2) (34–36). It is unclear if the recent upsurge in EV-D68 associated cases of pneumonia and flaccid paralysis-like illness are due to true changes in EV-D68 disease pathogenesis or improved molecular diagnostics (37). A previous study, using real-time reverse transcription polymerase chain reaction (qRT-PCR) assays, identified enteroviruses in 6% (515/8173) of patients that were hospitalised with pneumonia in South Africa (21). However, molecular characterisation of these viruses has not been done to determine whether some emerging lineages might be present in South Africa.

1.3.2.3 Echovirus 30

Enteroviruses are also a major cause of aseptic meningitis, and are usually associated with a favourable outcome. In temperate regions, seasonal aseptic meningitis usually peaks in summer and/or fall – similar to the seasonality observed in respiratory samples from South Africa (3,38). Many enterovirus genotypes circulate following an endemic pattern, but some genotypes such as echovirus 30 (E30) can cause widespread outbreaks. E30 is a member of enterovirus species B. The prototype strain of E30 (strain Bastianni) was isolated in 1958 (2).

E30 is one of the most common enterovirus genotypes responsible for sporadic as well as large-scale outbreaks of aseptic meningitis in many countries in Europe and North America (2,3). It is termed an epidemic virus since its circulation is highly variable from year to year (39). Phylogenetic studies with VP1 sequences indicate that E30 variants are continuously emerging and replacing the circulating

E30 strains, largely due to the error prone enterovirus RNA polymerase (31). Molecular epidemiology studies of E30 meningitis outbreaks show that point mutations give rise to substantial genetic diversity in the 1D gene encoding the VP1 capsid polypeptide (2). Twenty percent of codon sites in the 1D gene are subject to negative selection in most E30 population while a set of 10 codon positions were polymorphic and displayed signatures of relaxed negative selection (2). The substitution rate of E30 is 7.0 [95% high probability density, HPD 6.4-7.6] x 10^{-3} substitutions/site/year.

A common feature of E30 molecular epidemiology is progressive evolution of circulating lineages within one globally prevalent genotype (2,18). The dominant genotypes d-h (2) seem to be globally prevalent and all other lineages are seemingly extinct. E30 infection is usually cleared from the human host within a few days and, by analogy with another enterovirus infection - poliovirus, virus particles and virus infected cells are possibly cleared through the humoral response (15). The precise immune mechanisms involved in E30 infections have never been thoroughly investigated.

It has been reported that genotypes of E30 can become dominant for 3–4-year periods before they disappear from circulation (34), and they are therefore known as epidemic strains. E30 causes large outbreaks of aseptic meningitis that attract public health attention, and therefore there is a rich and extensive database of E30 VP1 sequences, from aseptic meningitis cases in the public domain. Despite the public health importance, relatively few complete genome sequences from sub-Saharan Africa are available for analysis (40). Only nineteen E30 full length genome sequences are currently available in GenBank, mainly isolated from aseptic meningitis cases with only two cases from respiratory illness cases from 2010, Thailand (https://www.ncbi.nlm.nih.gov/; accessed October 2018).

1.4 Rhinovirus species

1.4.1 Clinical manifestations and seasonality

Rhinoviruses are amongst the most commonly detected viruses in patients with upper respiratory tract infection and are considered to be associated, in most cases, with mild, self-limiting cold-like illnesses (41,42). The virus is usually mainly restricted to the upper respiratory airways and typically induces nasal congestion and rhinorrhoea, cough, sneezing, sore throat and malaise, with a spontaneous resolution within one to two weeks (43).

Rhinovirus infections are generally short-lived, and neutralizing antibodies that appear after viral clearance, have an essential role in protecting a person from reinfection. However, this humoral response seems to be genotype-specific with only little antibody cross-reaction (43). Rhinoviruses are detected throughout the year, with no evident seasonality (26). The large number of circulating genotypes supports the view that rhinoviruses do not circulate by waves or outbreaks of a given dominant genotype, which might explain the high frequency of reinfection during short periods (44).

1.4.2 Does rhinovirus contribute to the burden of severe respiratory illness?

Rhinoviruses have been identified from the nasopharynx of a large percent of healthy people without symptoms, and have not been considered as an important cause of acute respiratory hospitalisations in children (45). Recent advances in the development of sensitive molecular assays have considerably expanded the ability of laboratories to detect pathogens, especially viral agents. Together with the emergence of large pneumonia aetiology studies that include multiple viral causes, it has been shown that rhinovirus is not only associated with severe respiratory illness but also is identified frequently among asymptomatic patients (46,47).

Nonetheless, establishing a clinical association between pathogen detection and illness remains challenging since identification of a pathogen does not establish causality. A study performed in the USA on hospitalised children, estimated that

annual hospitalisation rates due to rhinovirus were 17.6/1000 in children less than 5 months followed by 6/1000 for children between 6-23 months and 2/1000 for children between the ages of 24-59 months. However when comparing children with or without a history of wheezing or asthma an increase in rhinovirus hospitalisations were observed 25.3/1000 vs. 3.1/1000, respectively (45).
In a previous study conducted in South Africa from 2012-2015, rhinoviruses were found to be significantly associated with mild and severe respiratory illness (21); whereas other studies could not establish any association (48–52). HIV infection is a known risk factor for increased severity of illness associated with the infection of several respiratory pathogens (7,53–55). In the South Africa study, 26% of enrolled patients with SARI were HIV infected and the HIV prevalence was 89% among patients aged 25-44 years (21). The significant association between rhinovirus detection and severe illness observed in this study but not in other studies raised the question whether the association was driven by the high HIV prevalence in South Africa.

1.5 Viral genomics, viral evolution and recombination events

Since enterovirus genotyping is based on partial VP1 sequencing, more extensive genome characterisation will assist in better understanding relationships between enterovirus genotypes since recombination frequently occurs in the P2 and P3 genomic regions (56). Complete genome sequencing is a powerful tool for pathogen characterization and are essential to describe circulating recombinant viruses (57).

In the last few years, complete genome sequences published in databases have grown exponentially along with the development of more sensitive and faster sequencing technology (58). Currently, next generation sequencing (NGS) is one of the most reliable and most popular technologies for this purpose. One of the strengths of NGS is ability to obtain detailed information about the pathogen in the sample in just one test in a relatively unbiased manner (57).

Viral evolution and recombination events can be studied using whole genome sequences. Enteroviruses evolve rapidly and even though most nucleotide substitutions are synonymous, amino acid changes do occur (59). Picornavirus RNA-dependent RNA polymerases are highly error prone and due to the absence of proof-reading activity, the substitution rate is high (1 mutation per 10^3 to 10^4 nucleotides) (60,61). Different parts of the enterovirus genome have distinct roles during the replication cycle and therefore may evolve differently and possibly exhibit remarkable independence during evolution (60). Mutation and recombination are the mechanistic alternatives for enterovirus evolution (60). The interaction between mutation (viral evolution) and recombination enables viruses to benefit from the advantages of having a few beneficial mutations without being affected by many deleterious mutations (40). Recombination occurs frequently in enteroviruses, thereby contributing to the genetic diversity and emergence of new genotypes. The first case of recombination in enteroviruses was originally reported in EV-B, specifically in polioviruses, and is since then recognised as a frequent and significant event occurring in enteroviruses (62). Recombination might occur within a specific genotype (intra-species) or between different genotypes of enteroviruses (inter-species)(40). To better understand enterovirus evolution, complete genome sequencing is important to identify inter and intra-typic recombination events, which might lead to emergence of strains with altered immune phenotypes, new tropisms, change in virulence or re-emergence of historic strains which may cause disease outbreaks and trigger serious public health concern (40,44,63).

The overall aim of this study was (i) to determine the epidemiology of enterovirus species by enterovirus type/ genotype and clinical syndrome during 2009-2014 and (ii) to determine the epidemiology of rhinovirus species and describe the fraction of rhinovirus infections attributable to illness, 2013-2015.

1.6 Aims and objectives

AIM 1: To determine the epidemiology and genotypes of enteroviruses associated with pneumonia in South Africa, 2009-2011

Objective 1: To determine the prevalence and seasonality of enterovirus in patients hospitalised with severe acute respiratory illness

Objective 2: To describe the clinical and epidemiological characteristics of patients hospitalised with severe acute respiratory illness testing positive for enterovirus

Objective 3: To determine the diversity of enterovirus genotypes present in patients hospitalised with severe acute respiratory illness by conducting traditional VP1 gene sequencing analysis

AIM 2: To compare enterovirus genotypes from patients with severe acute respiratory illness and influenza-like illness to those from asymptomatic controls in South Africa, 2012-2014

Objective 1: To determine the disease association, by HIV and age group, of enterovirus infection in patients with severe acute respiratory illness and influenza-like illness compared to asymptomatic controls

Objective 2: To determine the diversity of enterovirus genotypes present in patients with severe acute respiratory illness, influenza-like illness and asymptomatic controls by conducting traditional VP1 gene sequencing analysis

AIM 3: To perform genome-wide characterisation of most prevalent enterovirus, human echovirus 30, isolated from patients with severe acute respiratory illness in South Africa, 2013

Objective 1: To develop a viral enrichment sequence independent single primer amplification (SISPA) method to sequence the whole genome of enteroviruses

Objective 2: To conduct whole genome sequencing on echovirus 30 positive samples identified by traditional partial VP1 gene analysis in order to characterise viruses from respiratory illness cases across the complete genomes and to determine relatedness to strains reported elsewhere

AIM 4: To determine the epidemiology of rhinovirus species and to determine the attributable fraction of rhinovirus infection among HIV-infected and HIV-uninfected patients with mild and severe respiratory illness, South Africa, 2013-2015

Objective 1: To describe the clinical and epidemiological characteristics of patients hospitalised with severe acute respiratory illness and influenza-like illness testing positive for rhinovirus, compared to controls

Objective 2: To determine the attributable fraction of rhinovirus among patients hospitalised with severe acute respiratory illness and outpatients with milder influenza-like illness compared to asymptomatic controls.

Objective 3: To determine the adjusted prevalence corresponding to the prevalence of rhinovirus attributable to mild (influenza-like illness) and severe (severe acute respiratory illness), based on HIV status

Chapter 2

Materials and Methods

2.1 Description of the surveillance programmes and case definitions

2.1.1 Severe acute respiratory illness (SARI) surveillance

SARI surveillance is a prospective, hospital-based, surveillance programme conducted at two sites (Klerksdorp-Tshepong Hospital Complex, Klerksdorp, North West Province, and Edendale Hospital, Pietermaritzburg, KwaZulu-Natal Province) in South Africa (Figure 2.1).

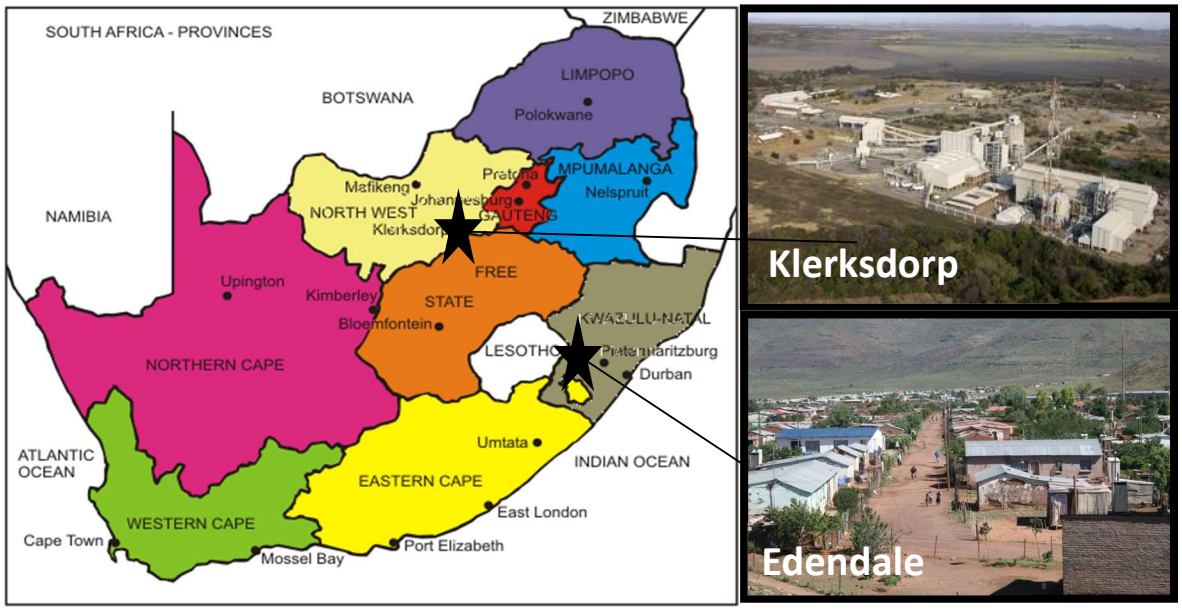


Figure 2.1: Map of provinces in South Africa – indicating the two surveillance sites (courtesy: Sibongile Walaza, Centre for Respiratory Diseases and Meningitis, NICD)

A case of SARI was defined as a patient who had symptom onset within 10 days prior to admission and who met age-specific clinical inclusion criteria as follows:

- A case in infants 2 days to less than 3 months of age included any patient with a diagnosis of suspected sepsis or physician-diagnosed acute LTRI, irrespective of signs and symptoms.
- A case in infants 3 months to less than 5 years of age included any patient with physician-diagnosed acute LRTI (e.g., bronchitis, bronchiolitis, and pneumonia) or pleural effusion.
- A case in individuals aged greater than and equal to 5 years of age included any patient with manifestation of acute LRTI with recorded temperature ≥38°C or history of fever and cough.

2.1.2 Severe chronic respiratory illness (SCRI) surveillance

A case of SCRI was defined as illness in a hospitalised person who had symptom onset greater than 10 days before admission and who met the age-specific clinical inclusion criteria described above (2.1.1).

2.1.3 Severe respiratory illness (SRI) surveillance

SRI cases include both SARI and SCRI cases. Therefore, a case of SRI was defined as a patient who was hospitalised (irrespective of symptom duration) and who met age-specific clinical inclusion criteria as follows:

- A case in infants 2 days to less than 3 months of age included any patient with a diagnosis of suspected sepsis or physician-diagnosed acute LTRI, irrespective of signs and symptoms.
- A case in infants 3 months to less than 5 years of age included any patient with physician-diagnosed acute LRTI (e.g., bronchitis, bronchiolitis, and pneumonia) or pleural effusion.

- A case in individuals aged greater than and equal to 5 years of age included any patient with manifestation of acute LRTI with recorded temperature ≥38°C or history of fever and cough.

2.1.4 Influenza-like illness (ILI) surveillance

ILI surveillance is a prospective, surveillance programme among patients presenting with ILI at two sites outpatient clinics (Jouberton Clinic, North West Province, and Edendale Gateway Clinic, KwaZulu-Natal Province) located in the same catchment area as the SARI and SCRI surveillance sites. An ILI case was defined as an outpatient of any age presenting with either a measured temperature greater than 38°C or, history of fever and cough with duration of symptom onset within 10 days.

2.1.5 Asymptomatic individuals

An asymptomatic individual or control was defined as persons of any age who sought medical care at one of the sentinel clinics for reasons other than fever, respiratory or gastrointestinal symptoms during the 14 days preceding the visit. Most controls visited the clinics for dental procedures, family planning, well-baby visits, voluntary HIV counselling and testing, or acute care for nonfebrile illnesses. The aim was to enrol one HIV-infected and one HIV-uninfected control every week in each clinic within each of the following age categories: 0–1, 2–4, 5–14, 15–54, and ≥55 years. Individuals were not followed-up to ensure they remained asymptomatic after enrolment.

2.2 Study procedures

Demographic and clinical characteristics were collected from all patients. Surveillance officers (registered/enrolled nurses) reviewed the clinical details of patients admitted to sentinel sites to establish whether cases met the surveillance

case definition/s. Identified cases were approached for inclusion in the sentinel surveillance. Consenting patients were asked to complete a questionnaire by interview on previous medical history and clinical data and to provide specimens.

Study staff completed case report forms for all enrolled controls, inpatients and outpatients. In addition, for inpatients, hospital records were reviewed to assess disease progression and outcome (i.e., discharge, transfer, or in-hospital death). Referral to hospital was recorded for all enrolled ILI cases. ILI cases that were referred to hospital were excluded from the analysis, as these patients did not adhere to the case definition of ILI anymore.

HIV results were obtained from a combination of two sources: (i) patient clinical records when available and (ii) for consenting patients, a dried blood spot was tested at the National Institute for Communicable Diseases (NICD) of the National Health Laboratory Service (NHLS), Johannesburg, South Africa. A HIV ELISA (Biorad Genscreen HIV ½ V2 EIA assay for screening and Biomerieux Vironostika HIV Ag/Ab EIA or DiaSorin Murex HIV Ag/Ab Combination EIA for confirmation) was used to test samples from patients aged ≥18 months, and PCR (COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Qualitative Test v2.0; Roche Molecular Systems, Inc., Branchburg, NJ) to test samples from children aged <18 months. Limited data were available on CD4+ T-cell counts on HIV positive patients.

2.3 Description of sample collection

Respiratory pathogen diagnosis depends on the collection of high-quality specimens, the rapid transport to the laboratory and appropriate storage before laboratory testing. Viruses are best detected in specimens containing infected cells and secretions.

Specimens should be collected as soon as possible, preferably during the first 72 hours (3 days) after onset of clinical symptoms in the acute phase of illness when

the greatest amount of viral shedding occurs, however specimens were processed even if collected greater than 3 days after the onset of symptoms.

For the optimal detection of respiratory pathogens, nasopharyngeal aspirates from children aged less than 5 years and combined nasopharyngeal-oropharyngeal swabs from persons aged greater than and equal to 5 years of age were collected from all enrolled participants. After collection, the specimens were placed in universal transport medium (Copan, Murruieta, California, USA), stored at 4–8°C, and transported within 72 hours of collection to the NICD for testing.

2.4 Laboratory methods and procedures

2.4.1 Total nucleic acid extraction

Total nucleic acid extraction for detection of enterovirus and rhinovirus was done using the MagNA Pure 96 instrument (Roche Diagnostics, Mannheim, Germany), which extracts 96 specimens per hour. Samples were prepared by adding 200 µl of each universal transport media in a processing cartridge. The extracted total nucleic acid (TNA) was eluted in 100 µl. The internal control was prepared by adding Poly (A) (final volume of 0.05 µg/ml) to the samples, in order to facilitate in the precipitation of RNA. The extraction was performed using the MagNA Pure 96 DNA and Viral Nucleic Acid Small Volume Kit together with the Viral Nucleic Acid Plasma Small Volume protocol.

Viral RNA for whole genome sequencing of enteroviruses was extracted from 200 µl viral cell culture supernatant using the total nucleic acid I kit and the Total Nucleic Acid Plasma_100_400 protocol on the MagNA Pure Compact instrument (Roche, Mannheim, Germany) according to manufacturer's instructions. The extracted total nucleic acid was eluted in 100 µl.

2.4.2 Laboratory diagnosis of enterovirus and rhinovirus

All respiratory specimens were tested for influenza A and B viruses, respiratory syncytial virus, parainfluenza virus types 1, 2 and 3, adenoviruses, human metapneumovirus, enteroviruses and rhinoviruses, using either a multiplex qRT-PCR assay (26) during 2013-2014 or the commercial Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) in 2015, with a Ct-value <37 regarded as a positive result.

For the multiplex qRT-PCR assay (26) on the Applied Biosystems 7500 Fast instrument (ThermoFisher, Waltham, Massachusetts, USA), complementary DNA (cDNA) was synthesized using the Transcriptor 1st Strand cDNA Kit (Roche Diagnostics), according to manufacturer's instructions. Each qRT-PCR reaction contained 2x LC Probe Master Mix (Roche, Mannheim, Germany), 1 µM of each primer, 0.5 µM of each probe, and 10 µl of cDNA reaction mixture as template, for a final volume of 30 µl. PCR cycles were initiated at 95°C for 15 minutes to activate thermostable DNA polymerase enzyme, followed by 45 cycles at 94°C for 15 seconds (denaturation), 60°C for 20 seconds (annealing), and 72°C for 10 seconds (elongation).

For the commercial Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) on the BioRad CFX96 instrument (Hercules, California, USA), each qRT-PCR reaction contained 5x Real-time One-step Buffer, 5x PR2 or PR3 primers/probes, 1U Real-time One-step enzyme and 8 µl of RNA as template, for a final volume of 25 µl. The initial step was reverse transcription at 50°C for 20 minutes followed by activation at 95°C for 15 minutes to activate thermostable DNA polymerase enzyme, followed by 45 cycles at 95°C for 10 seconds (denaturation), 60°C for 1 minute (annealing), and 72°C for 10 seconds (elongation).

The qRT-PCR used for enterovirus testing, is a sensitive method for the routine detection of all members of the enterovirus genus. This method has a weak cross-

reactivity to high-titre rhinovirus stocks, but not with rhinovirus-positive clinical samples (27).

2.4.3 Sanger sequencing of enterovirus VP1 gene region

A 375 base pair region targeting the 5' proximal part of VP1 region of enteroviruses were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of EV-68 (GenBank accession number AY426531). cDNA synthesis using random primers as well as oligo(dT) primers was performed with Transcriptor 1st Strand cDNA Kit (Roche Diagnostics, Mannheim Germany).

First round and nested PCR was performed as per Nix *et al*., 2006 (64). 5 µl of the cDNA reaction mixture was used in the first round PCR (final volume, 25 µl), consisting of 2x AmpliTaq Gold 360 PCR buffer (ThermoFisher, Waltham, Massachusetts, United States) and 10 pmol each of primers 224 (5'-GCI ATG YTI GGI ACI CAY RT-3') and 222 (5'-CIC CIG GIG GIA YRW ACA T-3'). PCR cycles were initiated at 95°C for 10 minutes to activate thermostable DNA polymerase enzyme, followed by 40 cycles at 95°C for 30 seconds (denaturation), 42°C for 30 seconds (annealing), and 72°C for 60 seconds (elongation). A final elongation step was performed at 72°C for 7 minutes.

2 µl of the first round PCR was added to a nested PCR, consisting of 2x AmpliTaq Gold 360 PCR buffer (ThermoFisher, Waltham, Massachusetts, United States) and 10 pmol each of primer AN89 (5'-CCA GCA CTG ACA GCA GYN GAR AYN GG-3') and AN88 (5'-TAC TGG ACC ACC TGG NGG NAY RWA CAT-3'), in a final volume of 25 µl. PCR cycles were initiated at 95°C for 10 minutes to activate Taq DNA polymerase enzyme, followed by 40 cycles at 95°C for 30 seconds (denaturation), 60°C for 30 seconds (annealing), and 72°C for 60 seconds (elongation). A final elongation step was performed at 72°C for 7 minutes.

5 µl of the nested PCR product (375 bp) were separated and visualized on a 1% agarose gel, stained with 2x GelStain Red (Biotium, Fremont, California, USA) and run for 1 hour at 120V using a 100 bp molecular weight marker (Roche, Mannheim, Germany) as a size reference. Slight variations in the sizes of the PCR products (350 to 400 bp) were observed due to VP1 gene length differences in the different genotypes.

Amplicons were purified using the ExoSAP-IT enzyme system (USB Corporation, Cleveland OH, USA). The exonuclease I and the Shrimp Alkaline Phosphatase were mixed in a 1:1 ratio and 3 µl of the mixture were added to 10 µl of amplicons. The mix were incubated at 37°C for 15 minutes followed by 80°C for 15 minutes.

Purified amplicons were sequenced using the Big Dye terminator version 3.1 cycle Sequencing Ready Reaction kit (Life Technologies, Foster City, USA) by using 10 pmol primers AN89 and AN88. 2 µl of the purified amplicon was used (final volume, 10 µl), consisting of 5x BigDye buffer, 5x BigDye mix and 10 pmol each of primer in the forward and reverse reaction, respectively. PCR cycles were initiated at 96°C for 0 seconds, followed by 25 cycles at 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 seconds.

After the sequencing reaction, the amplicons were cleaned using the BigDye Xterminator Purification kit (Life Technologies, Foster City, USA). 2 µl of 2.2 % SDS were added to 10 µl of cycle sequenced amplicons and incubated at 98°C for 5 minutes, 25°C for 10 minutes, and 15°C for 7 minutes, to reduce secondary structures. 10 µl X-terminator mix and 55 µl SAM solution were added to the SDS-treated amplicon. The reactions were shaked at 1800 rpm for 30 minutes and centrifuged at 1000 g for 2 minutes.

2.4.3.1 Assembly and phylogenetic analysis of enteroviruses partial VP1

Sequences were assembled using Sequencher® version 5 (Gene Codes Corporation, Michigan, USA). Sequence alignments were performed using MAFFT multiple sequence alignment program (65). The Kimura-2 parameter nucleotide substitution model determined as the optimal substitution model by using jModelTest was used for the Neighbor-joining (NJ) analysis (66,67) and the NJ trees were generated using MEGA 5.2 software (68,69).

2.4.4 Enterovirus isolation

Human epithelial type 2 (HEp-2) cells were used for expansion and passage of the E30 positive samples. Cells were cultured in DMEM (Dulbecco's Modified Eagle's Medium) containing 10% fetal bovine serum in an incubator at 37°C with 5% CO_2. Most cultures underwent two to four passages before being subjected for whole genome sequencing. Positive cultures were identified through characteristic cytopathic effect as seen under a light microscope. The virus was harvested by scraping infected cell layers and mixed with culture fluid by vortexing, and the suspension was subsequently centrifuged at 600 g for 10 minutes at 4°C to remove cellular debris.

2.4.5 Enterovirus genome sequencing

2.4.5.1 Viral enrichment and sequence independent single primer amplification (RACE-SISPA) of echovirus 30

2.4.5.1.1 Viral enrichment

The E30 culture supernatant fluids were filtered with a 0.22 micron filter, to remove bacteria and larger cellular components. To eliminate residual cellular DNA contaminants, the filtrate was treated with 0.1U/µl DNase I and incubated at

37°C for 1 hour. Viral RNA was extracted from 200 µl virus-enriched cell culture supernatant using the total nucleic acid isolation I kit on the MagNA Pure Compact instrument (Roche, Mannheim, Germany) according to manufacturer's instructions, and the RNA was eluted in 100 µl volumes. The eluates were treated again with 0.1U/µl DNase I.

A Nanodrop spectrophotometer (ThermoFisher, Waltham, Massachusetts, United States) was used to determine DNA concentrations and concentration >0.5 ng/µl was deemed sufficient for NGS. A bioanalyzer (Agilent technologies, California, United States) was used to assess the quality of the starting material and monitoring the size distribution after fragmentation and adapter ligation.

2.4.5.1.2 Rapid amplification of cDNA (RACE)

Poly-U RNA was generated by addition of 7 µl of the extracted RNA to a reaction mixture containing 10x NEBuffer 2 (New England Biolabs, Ipswich, Massachusetts, USA), 10 µM UTP, 2 units/µl poly-U polymerase and 40 units/µl RNase inhibitor (Promega, Madison, Wisconsin, USA) in a total volume of 20 µl. The reaction tubes were incubated at 37°C for 10 minutes and cooled to 4°C.

2.4.5.1.3 Sequence independent single primer amplification (SISPA)

SISPA was performed as per method describe by Djikeng *et al.,* 2008 (70) with slight modifications. First-strand cDNA synthesis using SISPA primer (10 µM), FR26RV-N (5'-GCC GGA GCT CTG CAG ATA TCN NNN NN-3'), as well as an anchored primer (1 µM), FR40RV_anchored A (5'-GCC GGA GCT CTG CAG ATA TCA AAA AAA AAA AAA VN-3'), was performed using Thermo Scientific Maxima H minus Reverse Transcriptase (ThermoFisher, Waltham, Massachusetts, United States). The Poly-U RNA (7.5 µl) was first denatured together with the SISPA (0.4 µM) and anchored primers (0.2 µM) as well as 2%

DMSO in a total volume of 12 µl for 5 minutes at 65°C, cooled to 4°C for 2 minutes and then moved to ice immediately.

The first-round cDNA synthesis reaction was completed by addition of the following reagents to the 12 µl denatured RNA/primer mix: 5 x RT Buffer, 200 units/µl Maxima H minus reverse transcriptase, 5 µM Venter SISPA primer (5'-GCC GGA GCT CTG CAG ATA TCG GCC ATT ATG GCC (Ribo-GGG)3') (70), dNTP mix (0.5 mM) (Promega, Madison, Wisconsin, USA), 40 units/µl RNase inhibitor (Promega, Madison, Wisconsin, USA) in a total volume of 20 µl. The reaction was subjected to the following conditions: 10 minutes at 25°C, 30 minutes at 50°C, 15 minutes at 65°C, 5 minutes at 85°C and cooled to 4°C.

Second-strand cDNA synthesis was done on each first strand cDNA product by adding 1 µl RNAse H (Ambion) and 1 µl Klenow large fragment DNA polymerase with 3'-5' exonuclease activity; and the reaction mix subjected to the following conditions: 60 minutes at 37°C, 15 minutes 75°C and cooled to 4°C. The double stranded cDNA products were purified using Qiagen MinElute PCR purification kit (Hilden, Germany) as per manufacturer's instructions. All samples were eluted with 12 µl elution buffer.

Random fragment amplification was done in a 50 µl reaction using 10 µl of double stranded cDNA as template together with 2x MyTaq Red mix, 10 µM primer FR20RV (5'-GCC GGA GCT CTG CAG ATA TC-3') (70) and nuclease-free water. The reaction was subjected to 98°C for 30 seconds, 40 cycles of 98°C for 10 seconds, 54°C for 20 seconds, 72°C for 45 seconds followed by a final extension of 72°C for 5 minutes and cooled to 4°C. A second shorter amplification was done. To each 50 µl sample reaction, the following was added: 12.5 µl 2x MyTaq Red mix, 1 µl primer FR20RV (10 µM) and 13.5 µl nuclease-free water. The whole mixture was subjected to 98°C for 30 seconds, 15 cycles of 98°C for 10 seconds, 54°C for 20 seconds, 72°C for 45 seconds followed by a final extension of 72°C for 5 minutes and cooled to 4°C.

The DNA products were purified using Qiagen MinElute PCR purification kit (Hilden, Germany) as per manufacturer's instructions. All samples were eluted with 12 µl elution buffer. The concentration of the DNA was quantified using Qubit® Fluorometer (Invitrogen, Carlsbad, California, United States) using manufacturer's instructions.

2.4.5.2 Next generation sequencing on the MiSeq platform

Viral genomes were sequenced with the Illumina MiSeq platform (Illumina, California, USA). Libraries were constructed using 1 ng of DNA with the Nextera XT DNA library Preparation kit (Illumina, California, USA). All clean-up steps were performed using Agencourt AMPure XP beads (Beckman Coulter, California, USA). The libraries were quantified and quality controlled using the Agilent High Sensitivity DNA Kit (Agilent technologies, California, USA). Cluster generation and sequencing were completed utilizing Illumina's standard protocol (71).

2.4.5.3 Data analysis: Trimming, mapping of the reads and assembly

Reads were de-multiplexed following the removal of tags and adaptors. Following importation in CLC Genomics Workbench (Qiagen, Hilden, Germany), reads were paired, trimmed and merged using the following parameters: Trim quality score limit = 0.01; Trim ambiguous nucleotides with maximum number of ambiguities = 1. After trimming, reads shorter than 50 nucleotides were discarded. For E30 the sequences were then mapped against the reference sequence (Human E30 strain Bastianni – GenBank accession number AF311938) using the following parameters: Mismatch cost = 10; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.5; Similarity fraction = 0.95. The consensus sequence was extracted and exported for further analysis.

2.4.6 Phylogenetic analysis of echovirus 30 whole genomes

Phylogenetic trees were constructed by the neighbour-joining method implemented in the MEGA 7.0.26 program (72) using the Kimura 2-parameter model. Regions containing alignment gaps were ignored in the analysis. The branch lengths of the dendrogram were determined from the topologies of the trees and were obtained by majority rule consensus among 1000 bootstrap replicates. Bootstrap values greater than 75% were considered statistically significant for grouping. Each gene segment was analysed with Basic Local Alignment Search Tool (BLAST) to confirm that no inter-typic recombination occurred (73). Similarity plot and boot scanning analyses were performed using the RDP4 software and DataMonkey (74–76). For similarity plot analyses, a 200-nucleotide window was moved in 20-nucleotide steps, and boot-scanning analyses were run with the neighbor-joining method. GARD (Genetic Algorithm for Recombination Detection) is a method to screen a multiple sequence analysis for the presence of recombination and is extremely useful as a pre-processing step for selection inference. Because recombinant sequences cannot be adequately described with a single phylogenetic history, selection inference on recombinant data often leads to a significant increase in false positives. GARD alleviates this concern by comprehensively screening an alignment for recombination breakpoints and inferring a unique phylogenetic history for each detected recombination block (76).

2.5 Statistical data analysis

Unconditional exact logistic regression was used to assess the association of enterovirus species with disease severity among patients with mild (ILI) or severe illness (SARI) using asymptomatic individuals as control group. Exact logistic regression was used to account for the fact that no EV-D species was detected among controls. EV-B was used as the comparison group as it was the species most frequently detected.

In addition, unconditional penalized logistic regression were used to estimate the attributable fraction (AF) of rhinovirus-associated hospitalisation and outpatient consultation by comparing the rhinovirus detection rate (observed) among ILI, SARI and SCRI cases to those of controls. Penalized logistic regression was selected to account for separation problems (i.e. no observations in one of the comparison cells) (77). The analysis was implemented among any inpatient with SRI (i.e., patients with SARI or SCRI combined). The AF was estimated from the odds ratio (OR) obtained from the logistic regression models as follows:

$$AF = \frac{OR - 1}{OR} * 100$$

Subsequently, the rhinovirus percentage positive associated with illness among patients with ILI, SARI, SCRI and SRI ($Prev_{Illness}$) from the observed percentage positive ($Prev_{Observed}$) was estimated as follows: $Prev_{Illness} = Prev_{Observed} * AF$
The analysis was implemented overall, and stratified by HIV serostatus within the following age categories: <1 year, 1-4 years, 5-24 years, 25-44 years, 45-64 years, ≥65 years as well as <5 years and ≥5 years. These age groups were selected for consistency and comparability with previous studies from South Africa.

All estimates were adjusted for co-infections with the other respiratory viruses investigated in this study and any underlying medical condition (treated as dichotomous variable indicating presence or absence of any evaluated condition) as they were significantly associated with the outcome. This was obtained by including covariates for any underlying medical condition and positivity for the individual viruses investigated in the study in the multivariable penalized logistic regression model.

Underlying medical conditions were asplenia, including asplenia or sickle cell anemia; chronic illness, including chronic lung, renal, liver, or cardiac disease, diabetes mellitus and asthma; other immunocompromising conditions (excluding HIV), including organ transplant, primary immunodeficiency, immunotherapy, and malignancy; neurologic disorders; burns; obesity; malnutrition; and

prematurity. Prematurity and malnutrition were included as underlying medical conditions only in the analyses implemented among children aged <1, 1-4 and <5 years. HIV infection and age (i.e., <1, 1–4, 5–24, 25–44, 45–64, ≥65) were also included as covariates in the non HIV-stratified overall model and the models implemented among children aged <5 (<1 and 1-4 years) and person aged ≥5 years (5-24, 25-44, 45-64 ≥65 years).

The Chi-squared or Fisher's exact test were used for comparison of categorical variables. Significant associations were considered at $p<0.05$. The statistical analysis was -implemented using STATA version 14.1 (StataCorp, College Station, Texas, USA).

Chapter 3

Enterovirus D68 and Other Enterovirus Genotypes Identified in South African Patients with Severe Acute Respiratory Illness, 2009-2011

Reference:

Hellferscee O, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, Cohen AL, Pretorius M, Cohen C, Madhi SA, Venter M. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza and Other Respiratory Viruses. 2017 May; 11(3):211-219

3.1 Introduction

Several respiratory viruses, including enterovirus have been associated with SARI in South Africa (21). Although the majority of enterovirus infections are subclinical; they can cause a variety of acute clinical syndromes including upper and lower respiratory tract illness (24).

A previous study identified enteroviruses in 6% (515/8173) of patients that were hospitalised with pneumonia in South Africa (26). However molecular characterisation of these viruses has not been done to determine whether some emerging clades might be present in South Africa. The contribution of specific clades to severe disease in light of the high HIV sero-prevalence in the country, were not investigated.

The aim was to describe the epidemiology of enteroviruses and determine which genotypes were circulating among South African patients hospitalised with SARI during 2009-2011.

3.2 Materials and Methods

3.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI surveillance programme (described in section 2.1).

3.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients as described in section 2.2 and 2.3.

Specimens were tested for the presence of 10 respiratory viruses (enteroviruses, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (26) (section 2.4.1-2.4.2).

The study period included samples collected from January 2009 to December 2011 (n=10,260). Due to challenges with availability of reagents, we did not test for adenovirus from August–October 2011. For this period, retrospective testing was done on 30% of randomly selected samples and the prevalence extrapolated. Enterovirus positive specimens from each year (2009: n=75/315, 2010: n=48/200 and 2011: n=77/327) were randomly selected for genotyping and molecular characterization to obtain 24% of all enterovirus positive samples per study year.

3.2.3 Amplification and sequencing of enteroviruses partial VP1

The 5' proximal part of VP1 region were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (see methods section 2.4.3).

3.2.4 Statistical analysis

Differences in characteristics of enterovirus genotypes were assessed using the Fisher's exact test. P-values <0.05 were considered to be statistically significant. Analysis was performed using STATA 13 (Stata Corporation, Texas USA).

3.3 Results

3.3.1 Clinical and epidemiological characteristics of patients hospitalised with enterovirus-associated SARI

During the study period, enterovirus was detected in 8.2% (842/10,260) of SARI cases. For observations with complete data, as displayed in Table 3.1, the majority of enterovirus -positive patients (86.5%, 726/839) were <5 years of age. Enteroviruses were detected throughout the study, although the detection rate was lower in 2010 (5.8%) than in 2009 (10.1%) and 2011 (8.9%, p=0.047). Sixty-three percent (501/795) of enterovirus cases had co-infections with other respiratory viruses and bacteria. Respiratory syncytial virus and adenovirus (21.5% and 30.6%, respectively) were the most frequently detected viral co-infections with enterovirus. The majority of patients (70.0%, 582/831) had a recorded fever and 25% (211/829) of cases had symptoms of ≥3 days prior to admission. Almost all enterovirus associated SARI cases (93.4%, 781/836) had antibiotics prescribed on admission and the majority (74.5%, 618/829) were hospitalised for less than three days.

The in-hospital case fatality rate was 2.2% (19/840). However, these deaths could not be attributed solely to enterovirus infection since these patients were also co-infected with various other pathogens: 15.8% (3/19) *Mycobacterium tuberculosis*;

11.8% (2/17) adenovirus; 10.5% (2/19) human metapneumovirus; 15.8% (3/19) parainfluenza virus 2; 10.5% (2/19) parainfluenza virus 3; 15.8% (3/19) respiratory syncytial virus and 5.3% (1/19) *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Streptococcus pneumoniae* (in blood samples).

HIV status was known for 64% (541/842) of cases, among whom the prevalence of HIV was 19.8% (107/541). Close to 62% (8/13) of the cases with fatal outcome was HIV positive. The HIV status among all SARI patients were known for 79% (8120/10260) of cases, among whom the prevalence of HIV was 49.52% (4021/8120), and the case fatality rate 5.6% (577/10260).

3.3.2 Molecular epidemiology of enterovirus species A to D

In this study 48 (24%) from the 200 randomly selected enterovirus positive viruses could be genotyped of which 3 grouped with the rhinoviruses; these were excluded from further analysis but are indicated as reference strains on the phylogenetic tree (Figure 3.1). EV-B (44%, 20/45) and EV-D (22%, 10/45) species were more prevalent than EV-A (16%, 7/45) and EV-C (18%, 8/45) with bootstrap support of 95-99% for clusters (Figure 3.1).

Enteroviruses were detected throughout 2009-2011, with no evident seasonality, although a higher EV-D68 (only EV-D species detected) activity was seen in 2011 (Figure 3.2A) whereas enterovirus species B, C and D co-circulated during 2009-2011 (Figure 3.2B) and EV-A was detected only in 2009 and 2010.

The majority of EV-A (86%, 6/7) and EV-B (55.0%, 11/20) strains were detected in 2010, while EV-C (62.5%, 5/8) and EV-D (70.0%, 7/10) were more commonly detected in 2011 (Table 3.1). The majority of patients infected with EV-B (80%, 16/20) and EV-D (80%, 8/10) had a recorded fever. A high proportion of viral co-infections (other than enterovirus) were found among EV-B (62.5%, 10/16) and EV-D (50%, 5/10) positive patients. All HIV-infected patients in this subset tested, were infected with EV-B-strains (Table 3.1, Figure 3.1).

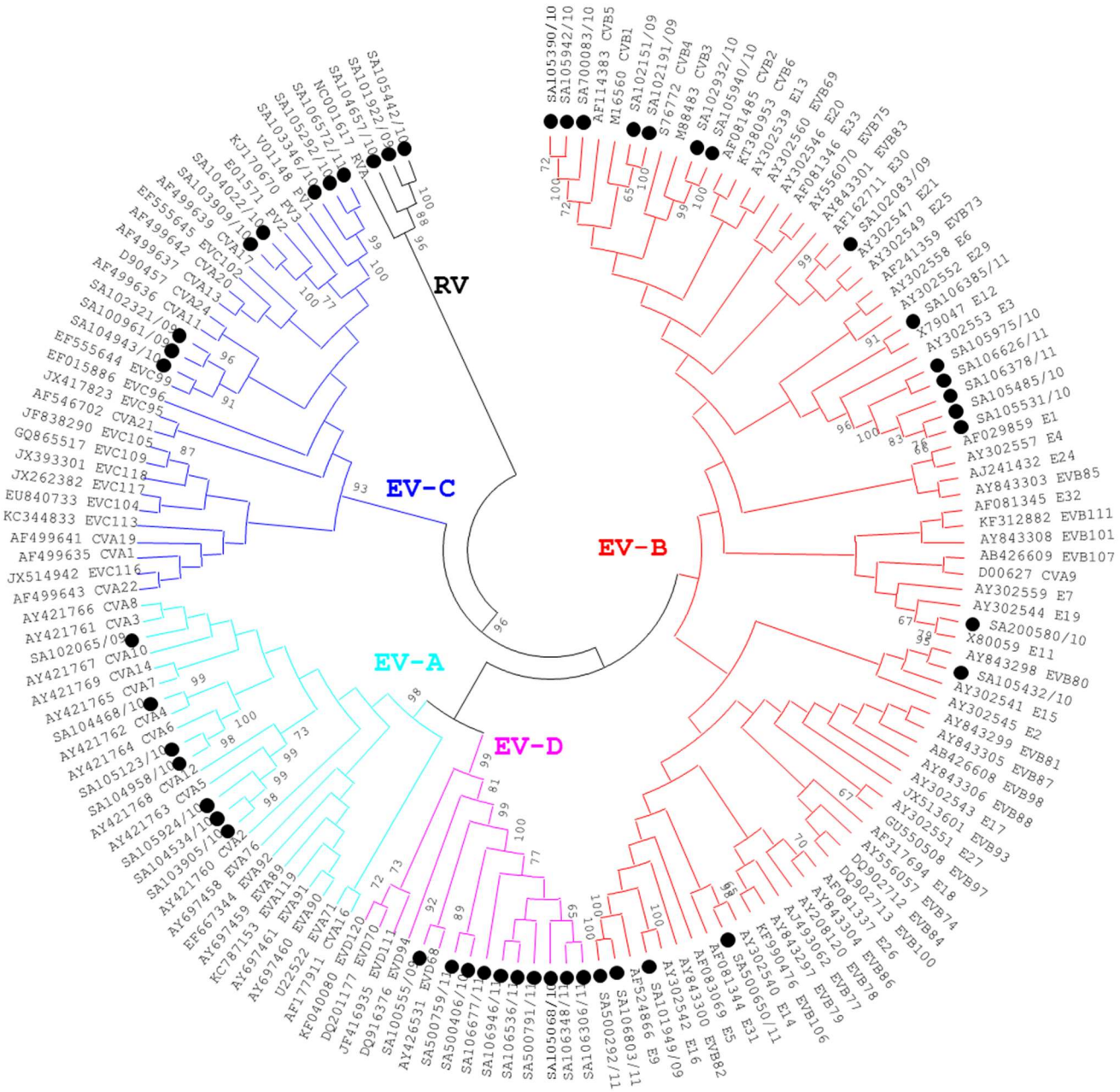
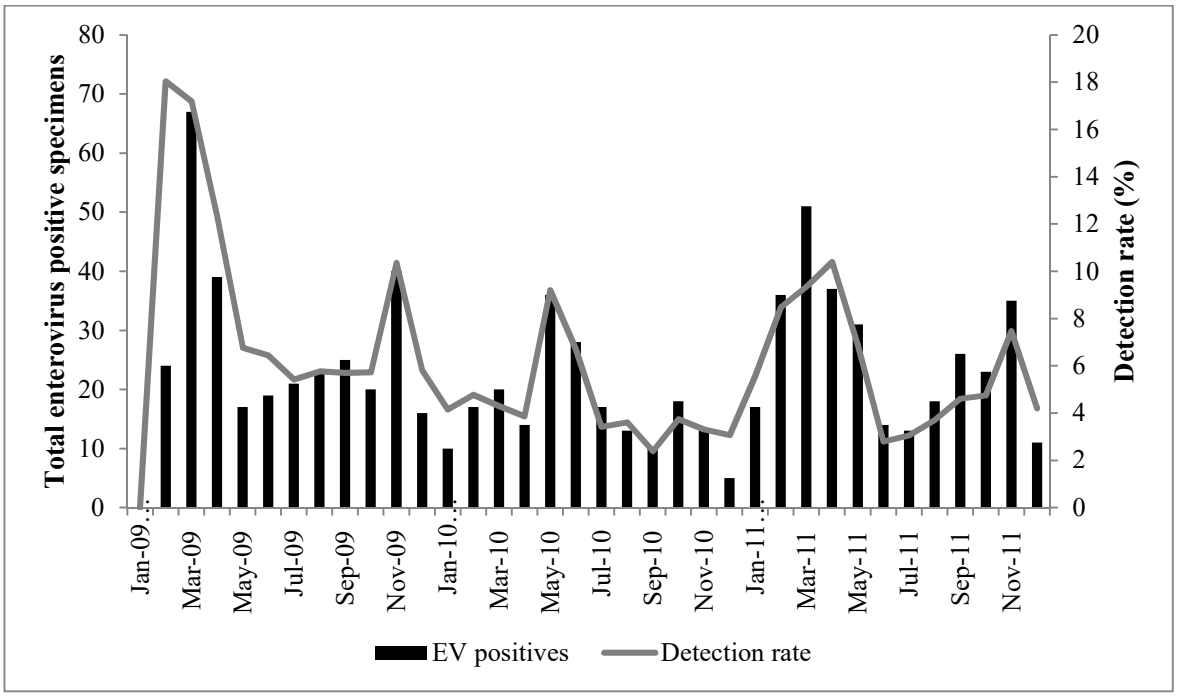


Figure 3.1: Molecular characterization of enterovirus clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implemented in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Black dots = South African samples 2009-2011.

A



B

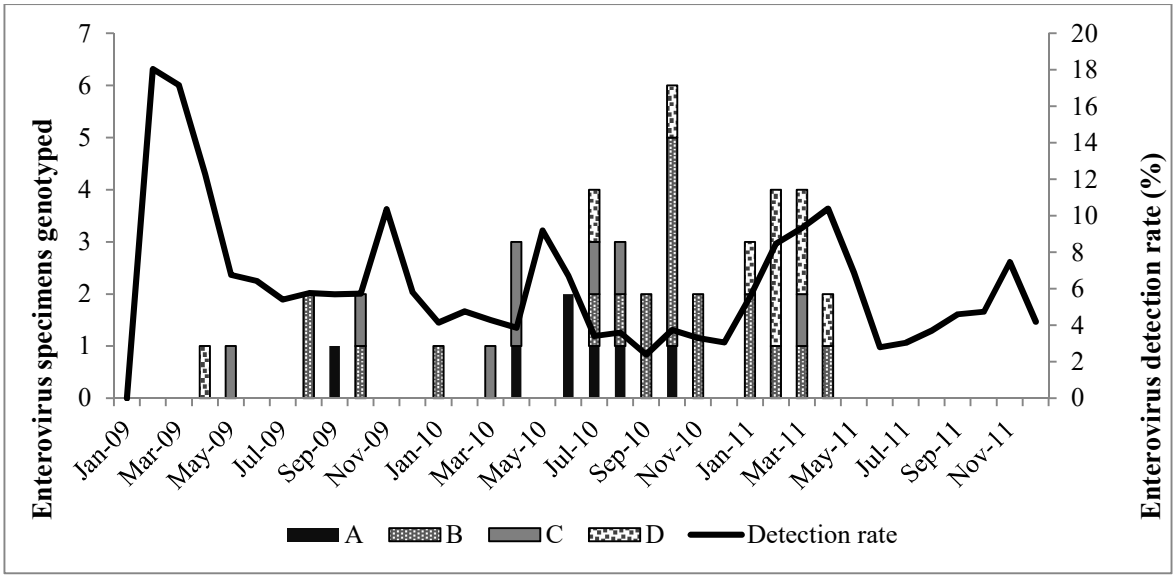


Figure 3.2: Epidemiologic graphs showing: (A) Distribution of enterovirus in South Africa during 2009-2011 and (B) Distribution of enterovirus species detected in South Africa during 2009-2011

3.3.3 Enterovirus genotypes and circulation of EV-D68 clades during 2009-2011

Seventeen distinct enterovirus genotypes were identified of which EV-D68 (22%, 10/45) and echovirus 3 (E3) (11%, 5/45) were the most frequently detected. Other genotypes detected include coxsackie virus A (3, 4, 5, 6), coxsackie virus B (1, 3, 5), echovirus (9, 11, 12, 14, 16, 30, 80, 99) and poliovirus (1, 2) (Figure 3.1).

Here EV-D68 strains are grouped according to phylogenetic clades A, B and C. The majority of EV-D68 strains (90%, 9/10) are in clade B. Previously reported South African EV-D68 strains from 2000-2001 (7) clusters in clades C and A (Figure 3.3). The South African clade B strains clustered with strains from respiratory specimens from Japan, 2010 and the Netherlands, 2009-2010 (8/10) and from the Philippines, 2008 and 2011 (1/10). The South African clade A strain clustered with strains from respiratory specimens from Italy, 2008, the Netherlands, 2008, the Philippines, 2011 and Kenya, 2010.

The designated clades B1 and B3, described recently were based on more complete gene sequence data (~930 base pairs for complete VP) (9, 10) and therefore the phylogenetic tree (based on 300 base pair VP1 fragment) did not have good bootstrap support for clade B1 and B3. Demographic and clinical characteristics of patients in whom EV-D68 was isolated are described in Table 3.2. Ages ranged from 3 to 54 months and respiratory symptoms included cough, chest indrawing, breathing difficulty and tachycardia.

No amino acid changes in the VP1 region that could lead to increased virulence (21) were observed in the South African samples (results not shown).

Table 3.1: Clinical and epidemiological characteristics of patients hospitalised with enterovirus associated severe acute respiratory illness at surveillance sites, South Africa, 2009-2011

Variables	All SARI samples n/N (%)	All enterovirus positive samples n/N (%)	Enterovirus genotyped n/N (%)	Species A n/N (%)	Species B n/N (%)	Species C n/N (%)	Species D n/N (%)	p value
Demographic and clinical characteristics								
Age (in years)								0.836
<1	4263/13598 (31.4)	377/839 (44.9)	23/45 (51.1)	3/7 (42.9)	10/20 (50)	5/8 (62.5)	5/10 (50)	
1-4	2411/13598 (17.7)	349/839 (41.6)	19/45 (42.2)	4/7 (57.1)	8/20 (40)	2/8 (25)	5/10 (50)	
≥5	6924/13598 (50.9)	113/839 (13.5)	3/45 (6.7)	0/7 (0)	2/20 (10)	1/8 (12.5)	0/10 (0)	
Male Sex	6640/13602 (48.8)	489/839 (58.2)	21/45 (46.7)	4/7 (57.1)	9/20 (42.9)	3/8 (37.5)	5/10 (50)	
Year								0.047
2009	3679/13656 (26.9)	315/3128 (10.1)	8/45 (17.8)	1/7 (14.3)	4/20 (20)	2/8 (25)	1/10 (10)	
2010	4613/13656 (33.8)	200/3460 (5.8)	24/45 (53.3)	6/7 (85.7)	11/20 (55)	5/8 (62.5)	2/10 (20)	
2011	5364/13656 (39.3)	327/3672 (8.9)	13/45 (28.9)	0/7 (0)	5/20 (25)	1/8 (12.5)	7/10 (70)	
Symptoms ≥3 days prior to admission	6867/13421 (51.2)	211/829 (25.5)	9/45 (20)	1/7 (15.3)	7/20 (35)	1/8 (12.5)	0/10 (0)	0.125
Fever (≥38˚C)	10755/13547 (78.4)	582/831 (70)	31/45 (68.9)	3/7 (42.9)	16/20 (80)	4/8 (50)	8/10 (80)	0.158

Oxygen therapy	4491/13508 (33.3)	236/831 (28.4)	15/45 (33.3)	2/7 (28.6)	6/20 (30)	1/7 (12.5)	6/10 (60)	0.215
Antibiotics prescribed on admission	12903/13549 (95.2)	781/836 (93.4)	41/45 (91.1)	7/7 (100)	17/20 (85)	8/8 (100)	9/10 (90)	0.809
Hospitalisation (≥ 3 days)	9053/13482 (67.2)	362/832 (43.5)	18/45 (40.9)	1/7 (14.3)	11/20 (55)	2/8 (25)	4/10 (44.4)	0.236
In-hospital death	755/13526 (5.6)	19/840 (2.2)	2/45 (4.4)	0/7 (0)	1/20 (5)	1/8 (12.5)	0/10 (0)	0.606
Underlying medical conditions								
HIV	5264/10666 (49.4)	107/541 (19.8)	4/34 (11.8)	0/5 (0)	4/17 (23.5)	0/3 (0)	0/9 (0)	0.412
Prematurity[a]	178/6661 (2.6)	21/725 (2.9)	2/45 (4)	0/7 (0)	1/20 (5)	1/8 (12.5)	0/10 (0)	0.606
Other medical conditions[b]	1189/13588 (8.8)	62/838 (7.4)	4/45 (8.9)	0/7 (0)	3/20 (15)	1/8 (12.5)	0/10 (0)	0.586
Co infections								
Streptococcus pneumoniae (blood)	706/10242 (6.9)	23/505 (4.6)	2/31 (6.5)	0/7 (0)	2/14 (14.3)	0/4 (0)	0/9 (0)	0.729
Respiratory syncytial virus	1897/13413 (14.1)	181/842 (21.5)	5/45 (11.1)	0/7 (0)	2/20 (10)	1/8 (12.5)	2/10 (20)	0.773
Para-influenza 2 virus	148/13413 (1.1)	20/842 (2.4)	2/45 (4.4)	0/7 (0)	2/20 (10)	0/8 (0)	0/10 (0)	0.798
Influenza virus (A and B)	1168/13430 (8.7)	33/842 (3.9)	2/45 (4.4)	0/7 (0)	2/20 (10)	0/8 (0)	0/10 (0)	0.798
Adenovirus	1129/12228 (15.8)	236/772 (30.6)	9/39 (23.08)	1/6 (16.7)	6/16 (37.5)	0/7 (0)	2/10 (20)	0.246
Any viral coinfection	7052/12860 (54.8)	501/795 (63)	19/39 (48.7)	3/6 (50)	10/16 (62.5)	1/7 (14.3)	5/10 (50)	0.205

Abbreviations: ICU: Intensive Care Unit; HIV: human immunodeficiency virus.

[a] Prematurity was classified as birth before 37 weeks of gestation as reported on the road-to-health card.

[b] Other evaluated underlying medical conditions included: asplenia or sickle cell anaemia; chronic illness, including chronic lung, renal, liver, or cardiac disease and diabetes; other immunocompromising conditions (excluding HIV), including organ transplant, primary immunodeficiency, immunotherapy and malignancy; neurological disorders and burns.

Table 3.2: Demographic and clinical characteristics of ten patients in whom EV-D68 was identified in South Africa

Characteristics	SA500759/11	SA106306/11	SA105068/10	SA106946/11	SA106677/11	SA106348/11	SA500406/10	SA500791/11	SA106536/11	SA100555/09
Demographic characteristics										
Age (in months)	25	15	3	11	3	16	54	16	11	5
Sex	F	F	F	M	M	F	M	M	F	M
Year	2011	2011	2010	2011	2011	2011	2010	2011	2011	2009
Medical history										
Underlying illness	N	N	N	N	N	N	N	N	N	N
HIV infected	N	N	N	N	N	N	N	N	UNK	N
Clinical presentation and course										
Symptoms (days) prior to admission	1	1	1	1	0	1	2	1	1	0
Fever (≥38˚C)	Y	Y	Y	Y	N	Y	Y	Y	Y	UNK
Respiratory symptoms	C, B, S, L, A	C, B, CI, T	C, B, T, A, V	C, T	C, CI	C, B, T	C, B, CI, S	C, B, T	C, B, CI, S, T, A, V	C, B, CI, T

Abbreviations: M: Male; F: Female; N: No; UNK: Unknown; Y: Yes; C: Cough; S: Stridor; R: Rhinorrhoea; M: Malaise; F: Fatigue; A: Anorexia; L: Lethargy; B: Breathing difficulty; CI: Chest indrawn; T: Tachycardia; V: Vomiting

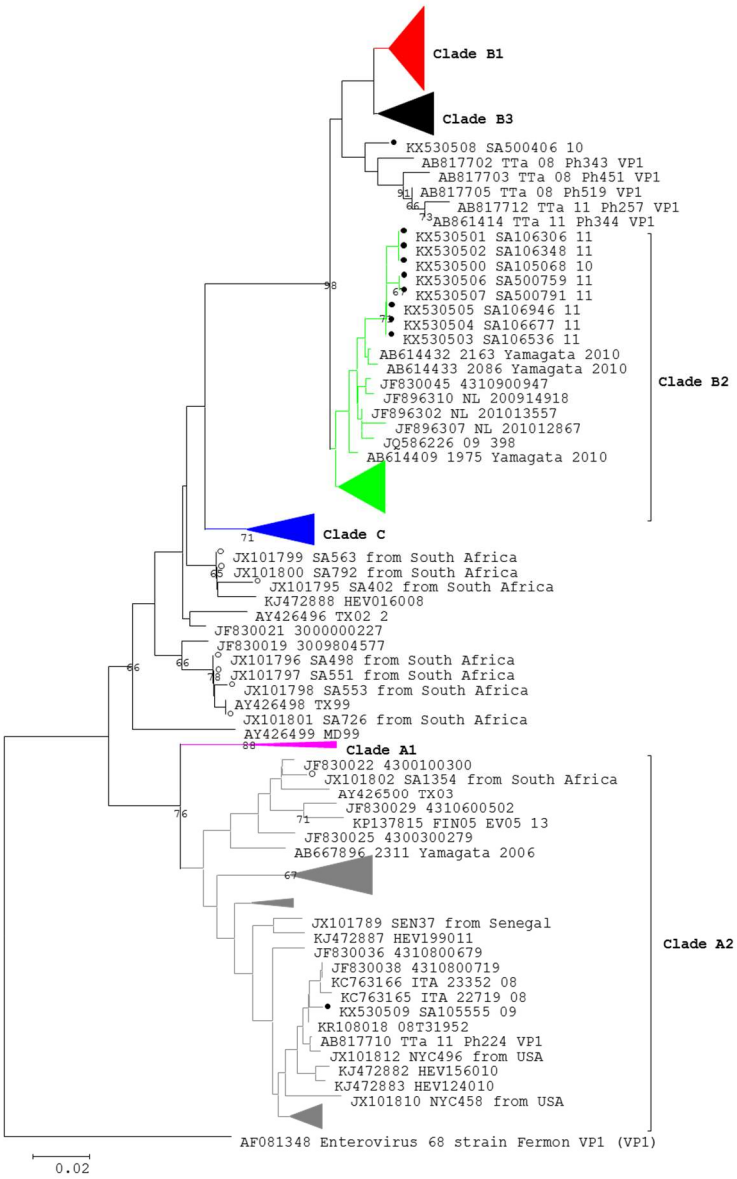


Figure 3.3: Molecular characterization of EV-D68 clinical samples based on phylogenetic analysis of nucleotide sequences of the partial VP1 genomic region. Trees were constructed using neighbor-joining methods as implicated in MEGA 5 software (http://www.megasoftware.net). Bootstrap values from 1000 replicates are shown on the nodes. Scale bar indicates number of nucleotide substitutions per site. Black dots = South African strains 2009-2011; Open dots = South African strains 2000-2001. Clades B1, B3, C A1 and partial B2 and A2 have been compressed to show comparison of South African strains to strains from other countries (country and year in blocks). Total number of sequences = 312 (representative strains from 2000-2016).

3.4 Discussion

During 2009-2011, enterovirus was detected in 8.2% of hospitalised SARI patients, which is within the range reported in other studies (3-25%) (39,78,79). Seventeen enterovirus genotypes were identified representing all four enterovirus species. EV-D68 was detected in 22% of enterovirus cases characterised (0.1%, 10/10260 amongst the total of hospitalised patients), which compares to frequencies reported in studies from The Netherlands (25%) and Germany (7.7%) (80,81). The EV-D68 strains that circulated in South Africa in 2009-2011 clustered in clades A and B of the EV-D species with the majority of strains grouping in clade B. All EV-D positive patients in this study were less than 5 years of age, which is similar to the typical younger age distribution reported for EV (78). A study to determine the worldwide emergence of EV-D68 included eight South African EV-D68 strains identified in hospitalised children over a one year period (2000-2001) (82). One of these South African strains (2001) was found within clade A, while the remaining seven (2000-2001) formed a cluster with sequences from Europe and the USA ancestral to clades B and C (82). More recent samples (2009-2011) were predominantly located in clade B of EV-D68.

It is speculated that the 2014 North American EV-D68 outbreak was likely due to modifications of the receptor interaction affecting the local cell tropism; the inability of previous established anti-EVD68 antibodies to detect the altered amino acid residues of the VP1gene (83).

A study from hospitalised patient with respiratory illness from New Zealand in 2010, indicates that EV-D68 may be associated with a more severe clinical presentation than other enterovirus strains (84). However, although the South African patients with EV-D68 detected in this study were all hospitalised, none were admitted to ICU or died. No recorded symptoms of neurological disease or paralysis for these patients was available, and no causative link between EV-D68 infection and neurological disease have been established to date (80). No EV-D positive samples were obtained from a South Africa study characterizing non-polio enteroviruses in stool and cerebrospinal fluid samples of patients symptomatic for gastroenteritis and acute

flaccid paralysis, respectively (personal communication, Wayne Howard, National Institute for Communicable Diseases). Higher EV-D68 activity in 2011 is consistent with another report (85). All polioviruses detected in this study were Sabin strains detected in the respiratory specimens of children that had recently been vaccinated.

This study has limitations that merit discussion. Only 24% of enterovirus positive samples selected for this study could be genotyped. The low success rate for genotyping may be due to low viral loads in clinical samples (Ct values >35 on original qRT-PCR assay) and subsequent inefficient nested amplification and sequencing of variable regions for phylogenetic analysis by degenerate PCR primers used for direct sequencing of PCR products. Furthermore, the study was limited to patients hospitalised with SARI and the presence of neurological manifestations was not specifically explored at time of participant enrolment. Identification of some enterovirus positives as rhinoviruses by molecular characterization could likely be due to cross reactivity in the 5` UTR due to sequence homology between enteroviruses and rhinoviruses.

Further investigation is needed to determine the disease association of enteroviruses and specifically EV-D68 infection with non-respiratory illness as well as full genome sequencing to determine possible recombination of EV-D68. Since this study did not include an asymptomatic control group, the study was limited in interpreting the attributing causality to enteroviruses in the presence of high frequency co-infections as well as determining concurrent prevalence in the healthy population.

In conclusion, the study showed that enteroviruses frequently contributed to SARI in South Africa. A high diversity in the enterovirus species that circulated in South Africa during 2009-2011, EV-D68 specifically circulated at high frequencies among children <5 years of age. Enteroviruses should be closely monitored to assess the emergence of highly pathogenic strains. There is no vaccine or antiviral available for non-polio enteroviruses. Understanding the contribution of enterovirus genotypes to severe illness will allow for informed decision making on potential candidate vaccines and development of therapeutic interventions.

Chapter 4

Enterovirus Genotypes among Patients with Severe Acute Respiratory Illness, Influenza-Like Illness and Asymptomatic Individuals in South Africa, 2012-2014

Reference:

Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, Madhi SA, du Plessis M, Cohen C, Treurnicht FK. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. Journal of Medical Virology 2017 Oct; 89(10):1759-1767

4.1 Introduction

Enteroviruses can cause symptoms similar to a mild cold, but have also been associated with severe respiratory infection requiring hospitalisation and may be fatal (86).

Though enteroviruses have been associated with SARI and ILI (21), these viruses have been less well characterized in South Africa. We identified high diversity among enteroviruses circulating in hospitalised South African patients (87), however it is unknown whether different enterovirus genotypes are associated with mild or severe respiratory illness in South Africa.

The aim was to characterize the enterovirus genotypes circulating among South African patients with SARI, ILI and asymptomatic individuals from June 2012 to

July 2014. In addition, the study assessed the association of different enterovirus species with mild (ILI) or severe (SARI) illness compared to asymptomatic individuals.

4.2 Materials and Methods

4.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI and ILI surveillance programmes as well as asymptomatic individuals from June 2012 through July 2014 (described in section 2.1).

4.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2 and 2.3.

Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (14) (methods section 2.4.1-2.4.2).

4.2.3 Amplification and sequencing of enteroviruses

The 5’ proximal part of VP1 region were amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (described in section 2.4.3).

The genetic diversity of each VP1 sequence was first determined by comparison with the reference strains in GenBank (US National Center for Biotechnology

Information, NCBI, http://www.ncbi.nlm.nih.gov/, accessed 01 July 2016) using BLAST and confirmed by phylogenetic analysis of the partial VP1 sequences. Sub analysis was similarly done for the three dominant enterovirus species in this study. Sequences of enterovirus partial VP1 genes generated in this study have been deposited in GenBank with the following accession numbers: KX940982-KX941096.

4.2.4 Statistical analysis

The Chi-squared or Fisher's exact test were used for comparison of categorical variables. Unconditional exact logistic regression was used to assess the association of enterovirus species with disease severity among patients with mild (ILI) or severe illness (SARI) using asymptomatic individuals as control group. Exact logistic regression was used to account for the fact that no EV-D species was detected among controls. EV-B was used as the comparison group as it was the species most frequently detected.

Significance was assessed for $p<0.05$. All models were adjusted for age, HIV serostatus and underlying medical conditions. The analysis was performed using STATA 14.1 (StataCorp, College Station, Texas, USA).

4.3 Results

4.3.1 Epidemiology of the study population

During the study period 1525 patients with SARI, 3088 with ILI and 1401 controls were enrolled, of which 1494 (97.60%), 3079 (99.7%) and 1367 (97.6%) were tested and had enterovirus results available, respectively. The detection rate for enterovirus was 5.8% (87/1494), 3.4% (103/3079) and 3.4% (46/1367) among SARI, ILI and controls, respectively (SARI/controls, p=0.002 and ILI/control, p=0.973).
In children <1 year of age, the majority (61.1%, 44/72) of enterovirus positive samples presented with SARI while in the following age groups: 1-4 years (43.1%,

50/116), 5-18 years (53.9%, 14/26) and >18 years (81.8%, 18/22) the majority of patients presented with ILI ($p<0.001$). Among enterovirus positive samples stratified in age groups above, HIV infection was not significant in any of the cases or controls (data not shown). Enteroviruses were detected throughout the year among SARI/ILI cases and controls (Figure 4.1 A-C), with peaks above 10% during February – March, August and November.

4.3.2 Phylogenetic comparison of enterovirus genotypes identified during June 2012 – July 2014

Of the 236 enterovirus -positive specimens detected during the study period among SARI and ILI cases and controls, 30.1% (71/236) were excluded from VP1 RT-PCR attempts, since the enterovirus load in these samples was considered low (Ct-value >35) and the attempt has a higher probability of failure. An additional 16.5% (39/236) of samples were excluded from VP1 RT-PCR and sequencing attempts, because the original specimen volume was insufficient for RNA extraction.

Of the enterovirus -positive samples that were sequenced (53.4%, 126/236), 7.9% (10/126) could not be typed. Possible reasons for typing failure included failed sequencing reactions, mixed enterovirus infections (co-infections) and rhinovirus that was detected on the initial enterovirus screen qRT-PCR assay. No cloning was attempted to rule out mixed infections. Rhinovirus were identified in 11.9% (15/126) samples, and were excluded from the enterovirus phylogenetic analysis. Therefore, only 101 enterovirus -positive samples were used for subsequent phylogenetic comparison.

Neighbor-joining phylogenetic comparison of enteroviruses detected in clinical samples from June 2012 to July 2014 to international reference sequences (downloaded from GenBank) indicated that EV-B (60.4%; 61/101) and EV-A (21.8%; 22/101) were more commonly identified than EV-C (10.9%; 11/101) and EV-D (6.9%; 7/101) (Figure 4.2). The majority of EV-A (40.9%, 9/22) and EV-D

(57.1%, 4/7) were detected among ILI cases and the majority of EV-B (55.7%, 34/61) and EV-C (54.6%, 6/11) were detected among SARI cases (p=0.03).

On multivariable logistic regression analysis using enterovirus species B as the reference group and controlling for age, HIV status and underlying illness, no association with disease severity (SARI or ILI compared to controls) was identified for any of the enterovirus species (Table 4.1).

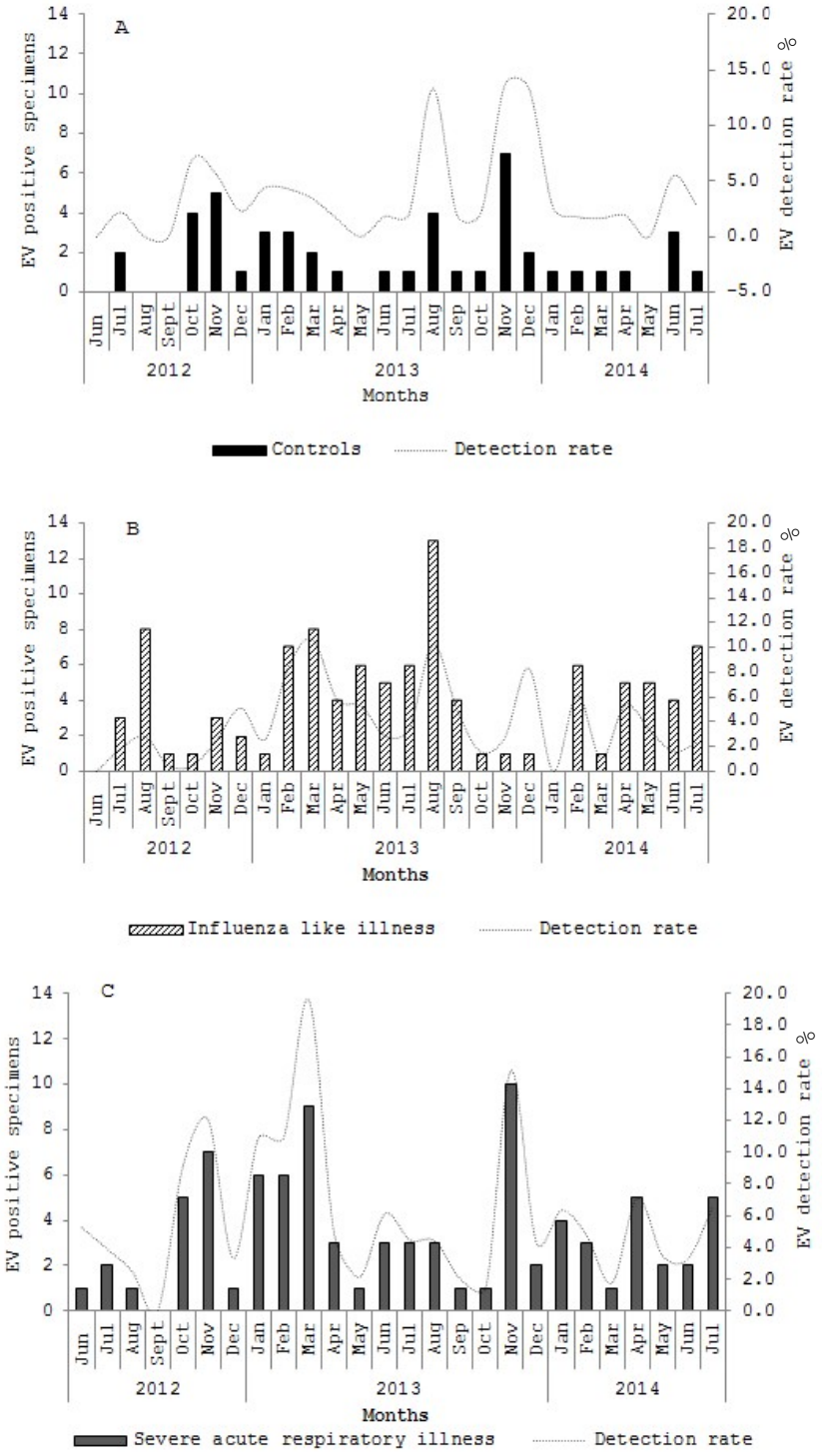


Figure 4.1: Monthly number and overall detection rate of enterovirus (EV)-positive cases among controls (A), influenza-like illness (B) and severe acute respiratory patients (C), Klerksdorp and Pietermaritzburg, South Africa, 2012-2014.

4.3.3 Enterovirus group and genotypes

Enterovirus genotypes could be identified in 59% (51/87), 33% (34/103) and 35% (16/46) of enterovirus-positive samples among SARI and ILI cases and controls, respectively. A total of 33 genotypes were identified, distributed among all four enterovirus species. The most prevalent genotype was E30 (9.9%, 10/101), followed by coxsackievirus B5 (7.9%, 8/101) and EV-D68 (6.9%, 7/101). All other genotypes identified in the study were detected in <5% of genotyped samples (Figure 4.3).

Coxsackievirus A14 (CVA14), echovirus 11 (E11), echovirus 16 (E16), echovirus 17 (E17), echovirus 19 (E19), echovirus 7 (E7), echovirus 9 (E9), poliovirus 1 (Sabin vaccine strain), poliovirus 3 (Sabin vaccine strain) were detected only among SARI cases and coxsackievirus A10 (CVA10), coxsackievirus A16 (CVA16), coxsackievirus B1 (CVB1), echovirus 32 (E32) and coxsackievirus A20 (CVA20) were detected only among ILI cases. Coxsackievirus B5, E30, echovirus 6 (E6) and EV-D68 were detected among SARI and ILI cases only (Figure 4.3).

EV-A, EV-B, EV-C and EV-D genotypes co-circulated in 2012-2014, while EV-D68 was detected sporadically in March-May in 2013 and 2014 only (Figure 4.4). The majority of enterovirus positive samples that could be genotyped, were from children aged <5 years (89.1%, 90/101) in the EV-B species (60.4%, 61/101) compared to the other enterovirus species, but were not statistically significant (p=0.08).

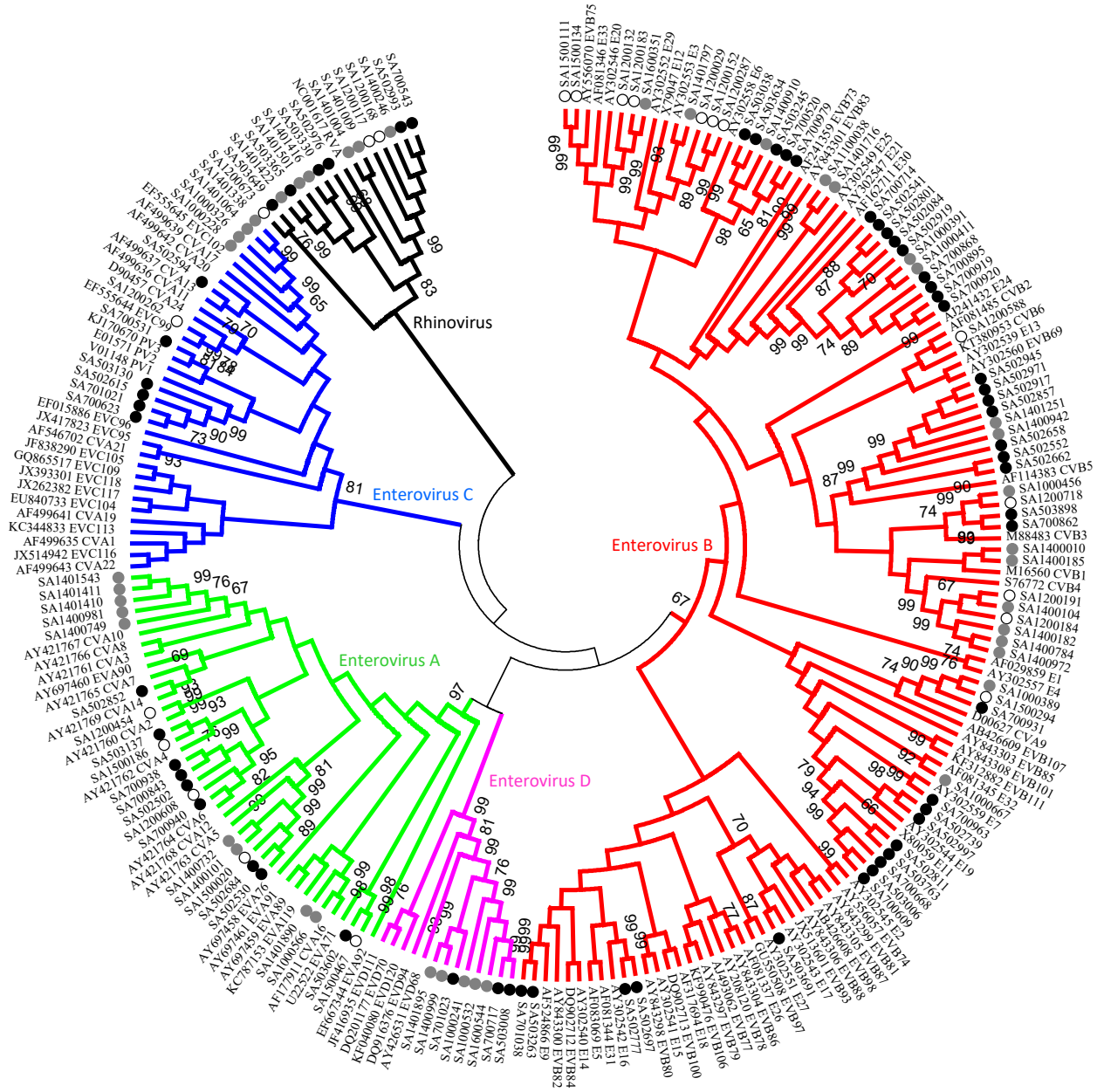


Figure 4.2: Phylogenetic analysis of enterovirus genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denotes genotypes identified in SARI patients, those with grey closed circles denotes genotypes identified in ILI patients while those with open circles denoted genotypes identifies in asymptomatic individuals. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree.

Table 4.1: Association of enterovirus species (EV-A to EV-D) with influenza-like illness (ILI) and severe acute respiratory illness (SARI) compared with asymptomatic controls, South Africa, 2009-2010 (bold indicate factors significant at p<0.05)

Factor	Prevalence			Multivariable analysis	
	Control	ILI	SARI	ILI	SARI
	n/N (%)	n/N (%)	n/N (%)	aOR* (95% CI)	aOR* (95% CI)
EV-A	5/16 (31)	9/34 (26)	8/51 (16)	1.16 (0.15-8.95)	1.43 (0.22-9.22)
EV-B	10/16 (63)	17/34 (50)	33/51 (66)	Reference	Reference
EV-C	1/16 (6)	4/34 (12)	6/51 (12)	0.41 (0.02-6.98)	0.68 (0.05-9.63)
EV-D	0 (0)	4/34 (12)	3/51 (6)	1.24 (0.11-∞)	0.53 (0.39-∞)

*Estimated odds ratio (aOR) adjusted for age, HIV and underlying medical condition

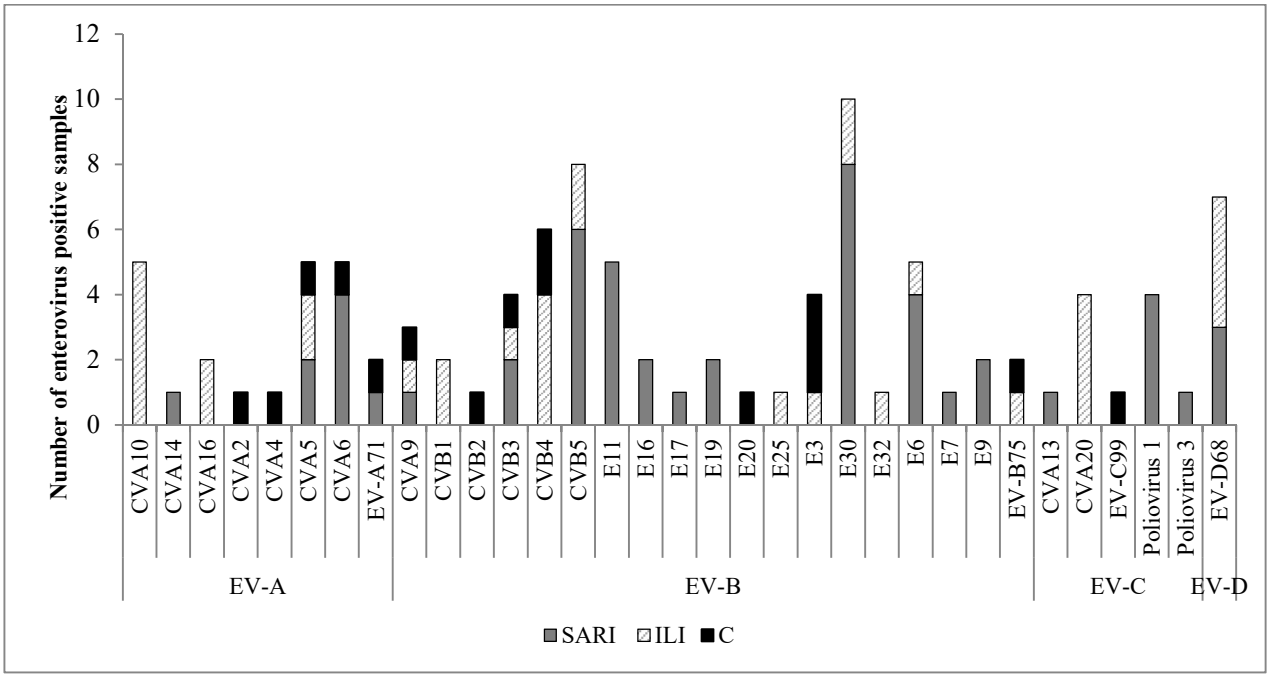


Figure 4.3: The genotypes and species (EV-A to EV-D) of enteroviruses identified in patients with severe acute respiratory illness (SARI), influenza-like illness (ILI) and asymptomatic individuals (C) from June 2012 to July 2014, South Africa

(CVA10=Coxsackie virus A10, CVA14=Coxsackie virus A14, CVA16=Coxsackie virus A16, CVA2=Coxsackie virus A2, CVA4=Coxsackie virus A4, CVA6=Coxsackie virus A6, EV-A71=Enterovirus A71, CVA9=Coxsackie virus A9, CVB1=Coxsackie virus B1, CVB2=Coxsackie virus B2, CVB3=Coxsackie virus B3, CVB4=Coxsackie virus B4, CVB5=Coxsackie virus B5, E11=Echovirus 11, E16=Echovirus 16, E17=Echovirus 17, E19=Echovirus 19, E20=Echovirus, E25=Echovirus 25, E3=Echovirus 3, E30=Echovirus 30, E32=Echovirus 32, E6=Echovirus 6, E7=Echovirus 7, E9=Echovirus 9, EV-B75, CVA13=Coxsackie virus A13, CVA20=Coxsackie virus A20, EV-C99=Enterovirus C99, EV-D68=Enterovirus D68)

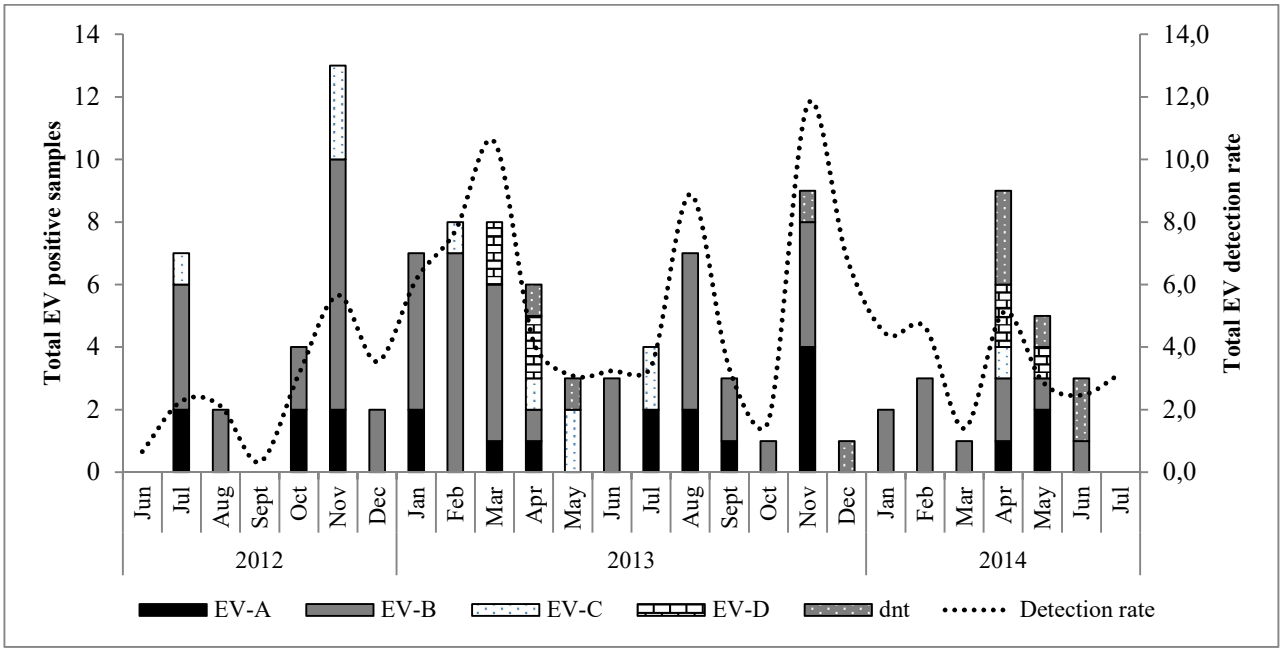


Figure 4.4: Distribution of enterovirus (EV) species (A-D) in South Africa during June 2012 to July 2014; dnt=did not genotype

4.3.4 Molecular evolution of echovirus 30, coxsackie virus B5 and EV-D68 genotypes

4.3.4.1 Echovirus 30

Phylogenetic analysis of E30 samples showed that all E30 samples from this study clustered with a single Australian strain from 2000 (GU232849) with 98% bootstrap support. Following the classification convention for E30 as described by Bailly *et al* (2009), this study propose to call this sub cluster genotype k (2). The sequences depicted in Figure 4.5 represent all known genotypes.

All E30 strains identified in this study displayed >98% nt homology and formed a distinct bi-phyletic cluster with 79% bootstrap support (Figure 5). In the bi-phyletic cluster, strains from Edendale SARI cases and Klerksdorp SARI and ILI cases are distinctly located in each of the sub clusters with 92% and 93% bootstrap support, respectively.
The E30 strains in this study differ from other E30 genotypes by mean nucleotide pairwise distance of 20.2% (15.6-27.5) and mean amino acid sequence distance of 6% (2.7-11.6) in amino acid sequence.

4.3.4.2 Coxsackie virus B5

All coxsackie virus B5 strains identified in this study (6 SARI samples and 2 ILI samples from Edendale) were categorized as genogroup C and formed a distinct sub cluster with a bootstrap value of 91% (Figure 4.6). The nucleotide sequences for coxsackie virus B5 strains were downloaded from GenBank and sequences depicted in Figure 6 represent all known genogroups. 6/8 samples (4 SARI and 2 ILI samples) were identical (nt homology 100%). All of the coxsackie virus B5 strains (identified in this study) displayed close genetic relationships with samples identified in surface water from Belarus, 2007 (GQ352398) and faeces from Germany, 2009 (HF948277).

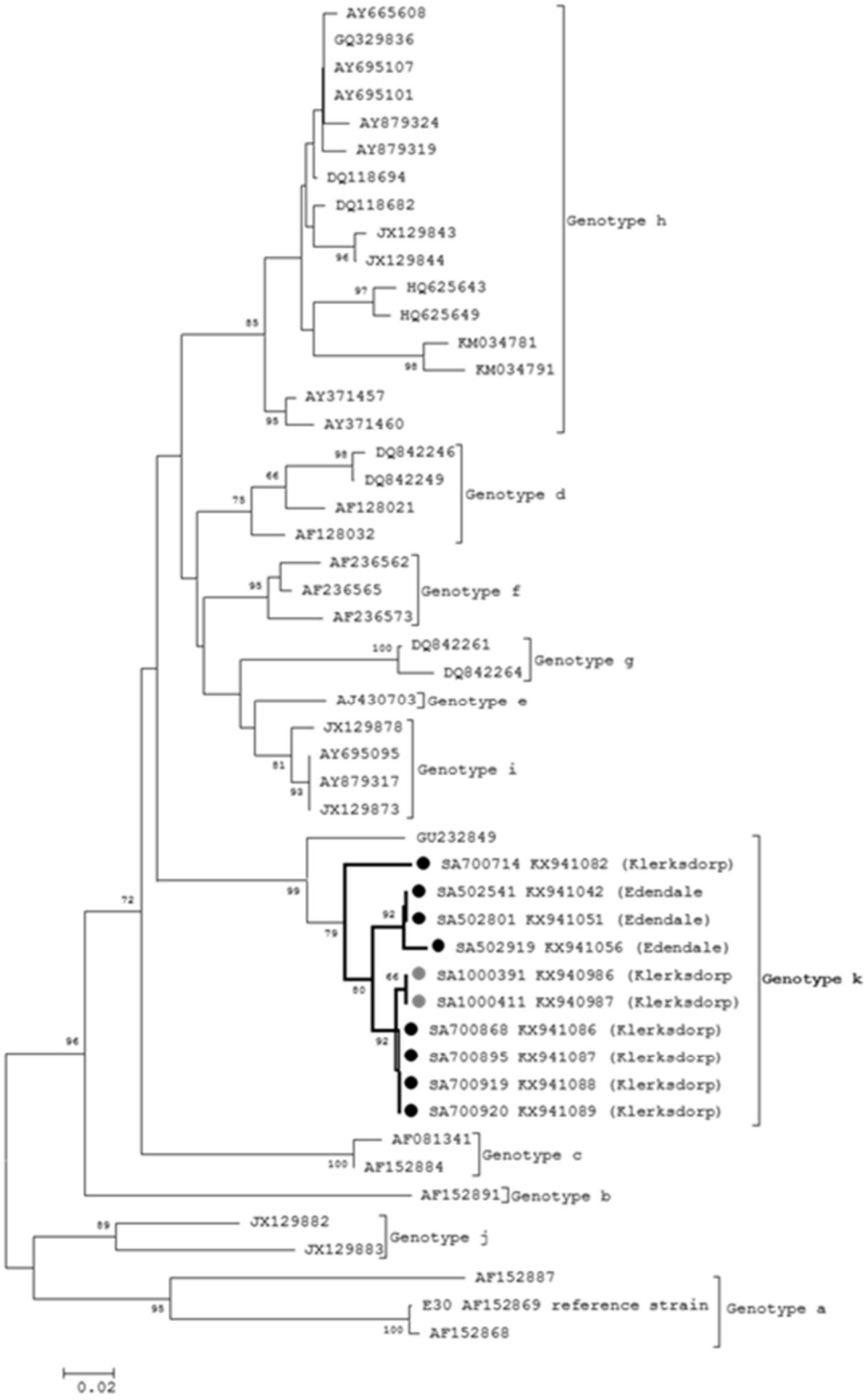


Figure 4.5: Phylogenetic analysis of echovirus 30 genotypes by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

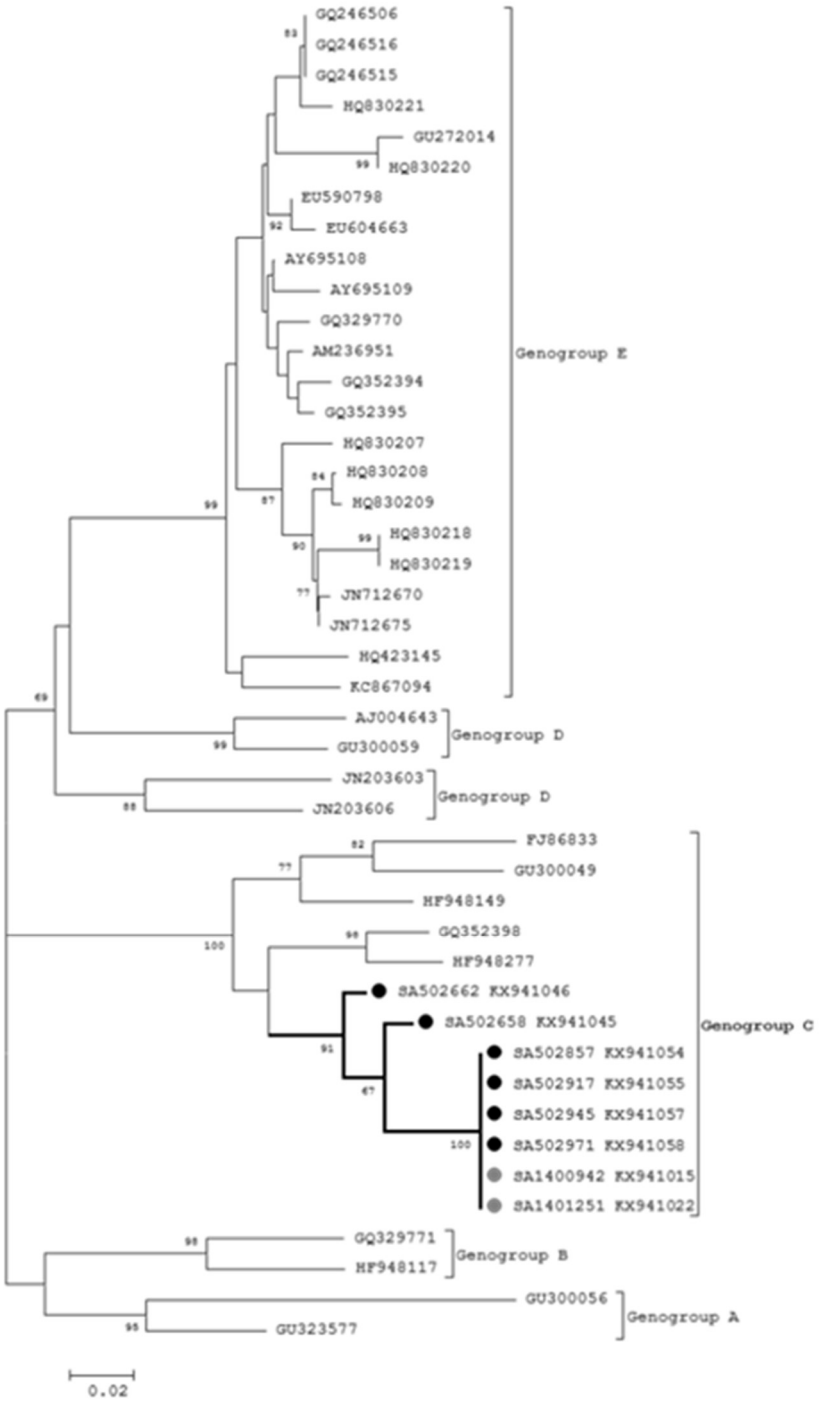


Figure 4.6: Phylogenetic analysis of coxsackie virus B5 genogroups by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

Sequences from the South African coxsackie virus B5 strains differed from the other coxsackie virus B5 genogroup C sequences by 13.4% in pairwise nucleotide sequence distance and 3% in pairwise amino acid sequence. Strains of coxsackie virus B in samples 502662 and 502658 which were not identical to the other South African sequences were sampled in 2012 and 2013, respectively.

4.3.4.3 EV-D68

The majority (71.4%, 5/7) of EV-D68 strains identified in 2013/2014 distinctly clustered in lineage B2 (89% bootstrap support) of which 3 were ILI cases and 2 were SARI cases, together with strains from one of the two co-circulating EV-D68 lineages that caused the large 2014 USA outbreak (Figure 4.7). The nucleotide sequences for EVD-68 strains were downloaded from GenBank and sequences depicted in Figure 4.7 represent all known lineages.
The lineage B EV-D68 strains in this study differed from other lineage B EV-D68 strains by 1.5% nucleotide and 1.4% amino acid sequence. These included 2 SARI cases (1 from Klerksdorp and 1 from Edendale) and 3 ILI cases (1 from Klerksdorp and 2 from Edendale).

The other South African EV-D68 strains (28.6%, 2/7) belonged to lineage A forming a unique cluster. The two strains from 2013 are from the same geographic location and clustered with 99% bootstrap support. The lineage A EV-D68 strains in this study differed from other lineage A EV-D68 strains by 12.8% nucleotide and 3.0% amino acid sequence. The 2012-2014 South African EV-D68 strains did not cluster with strains of EV-D68 previously identified in South African patients ((34) ,unpublished data).

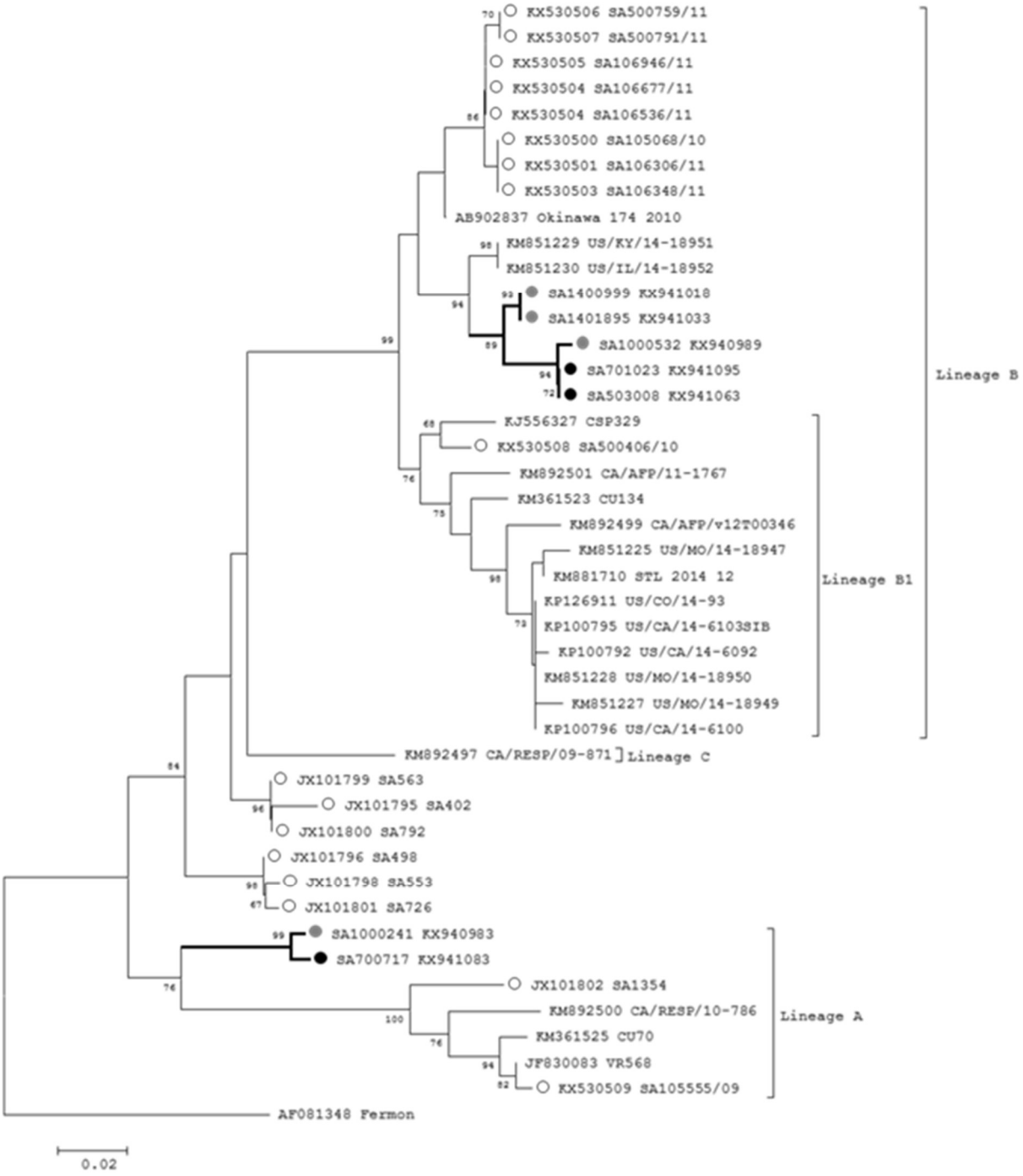


Figure 4.7: Phylogenetic analysis of EV-D68 lineages by neighbor-joining phylogenetic analysis of the partial VP1 region, South Africa, 2012-2014. Sequences with black closed circles denote genotypes identified in SARI patients and those with grey closed circles denote genotypes identified in ILI patients. Sequences with open circles are EV-D68 strains previously identified from South Africa. Bootstrap values (1000 replicates) are shown on the branches, with values <65% omitted from the tree. The scale indicates number of base substitutions per site.

4.4 Discussion

This study observed a high diversity of circulating enterovirus genotypes (a total of 33 genotypes) from enterovirus species A-D with high circulation rates for EV-B and EV-A compared to EV-C and EV-D. No disease association of enterovirus species with disease severity were determined.

From the systematic surveillance across all age groups, the majority of patients testing positive for enterovirus were < 5 years of age. This study did not observe any difference in disease association due to different enterovirus species. This could be firstly because of a lack of power (small sample numbers) and since enterovirus disease association cannot be conclusively ascribed since the analysis did not include co-infections with other pathogens. Epidemics of human enterovirus disease display a seasonal pattern, with infections more common in summer and early autumn in geographical regions with a temperate climate (3,38). Most clusters of EV-D68 showed an atypical late seasonality compared to other enteroviruses, with a peak in autumn, instead of summer (3). Outbreaks of enterovirus disease in autumn, early or late, are typical. Distribution of EV-A through –D in most temperate zones is from spring through autumn, with peak disease most years in late summer-autumn. This study showed similarities in the seasonal distribution of the enterovirus species detected in South Africa as enterovirus species A-D were detected throughout the year, but EV-D68 was sporadically detected in autumn (March-May) during 2013-2014.

E30, coxsackie virus B5 and EV-D68 were the genotypes most frequently identified in this study. Within a genotype, enteroviruses are deemed relatively uniform, as was observed in this study for E30 and coxsackie virus B5. However, some studies have suggested variation within a genotype (88), such as was observed for EV-D68 in this study.

E30 is the most common enterovirus strain responsible for sporadic as well as large-scale outbreaks of aseptic meningitis in many countries of Europe and North America (2). E30 causes large outbreaks of aseptic meningitis that attracts public

health attention and as a result there is a rich and large database of E30 VP1 sequences regarding aseptic meningitis, E30 has however not been widely associated with respiratory illness.

A common feature of E30 molecular epidemiology is the progression of circulating lineages within one prevalent genotype (2,18). The E30 strains identified in this study clustered together (designated genotype k) although no differences were observed between viruses from SARI compared to ILI cases. There is no report of a concurrent E30 aseptic meningitis outbreak at the same time the E30 combined nasopharyngeal and oropharyngeal specimen (NP/OP) positives were detected. It has been reported that genotypes of E30 can become dominant for 3-4-year periods before they disappear from circulation (88), and they are therefore known as epidemic strains. In this study, it seems that E30 genotype k became dominant and circulated in South Africa during 2012-2014. Phylogenetic studies with VP1 sequences indicate that E30 variants are continuously emerging and replacing the circulating E30 strains, largely due to the error prone enterovirus RNA polymerase (2).

Coxsackie virus B5 has been detected for over 50 years and, similar to E30 , sporadic cases of aseptic meningitis, as well as outbreaks, have been reported, remaining one of the most predominant reported enterovirus genotypes in a number of countries (89). All of the coxsackie virus B5 strains identified in this study formed a sub cluster in genogroup C. Phylogenetic clustering by year of study enrolment was observed for the South African strains. Different co-circulating strains of enterovirus are often observed in enterovirus disease outbreaks as enterovirus genomes vary with respect to time but may also vary by geographical distribution (90).

In 2014, the largest outbreak to date of severe respiratory illness associated with EV-D68 occurred in the USA, and almost all of the case patients were children (91). Coinciding with the 2014 USA outbreak there was an increased number of reports of children with sudden onset acute flaccid myelitis. Some recent South

African EV-D68 strains clustered with USA 2014 lineage B2 EV-D68 strains. The USA 2014 lineage B2 co-circulated with lineage B1 EV-D68 during the large 2014 USA outbreak. EV-D68 lineages B1 and B2 are both temporally associated with acute flaccid myelitis (92). The majority of EV-D68 strains identified in this study clustered in lineage B2 and the remaining samples clustered in lineage A, forming a unique cluster. The 2012-2014 South African samples did not cluster with strains previously identified in South Africa during 2009-2011.

The mechanisms for emergence of enterovirus-associated outbreaks are not known; however, a combination of virus-specific, population-level and other external factors are likely to be involved (3).

This study has some limitations. Previous studies observed a difference in the predominant enterovirus genotypes among neonates and older children (93) but this study was not powered for this analysis. Some rhinovirus sequences were detected in samples that tested positive for enterovirus using the in-house assay indicating that this enterovirus detection assay is not 100% specific to detect only enteroviruses. This study was not powered to describe the role of co-infections of other respiratory pathogens, and there for the role of enterovirus causing SARI and ILI could be overestimated.

In conclusion, this study showed that there was a high diversity in the VP1 sequences of the enterovirus species circulating in South Africa during 2012 to 2014 and most of these genotypes are associated with meningitis worldwide. Enterovirus was detected in outpatient and hospitalised patients with SARI but was also detected in controls. No disease association of enterovirus species with disease severity were found; however, some genotypes (E30, coxsackie virus B5, EV-D68) were more prevalent in symptomatic cases. Further studies are needed to determine if other factors such as viral load or host interactions play a role in enterovirus -associated disease as well as a robust spatiotemporal phylogenetic analysis based on complete VP1 sequences.

Chapter 5

Genome Analysis of Human Echovirus 30 Isolated from Patients with Acute Respiratory Illness in South Africa, 2013

5.1. Introduction

E30 was the most frequently identified enterovirus genotype in symptomatic cases (overall 9.9%, 10/101) in a previous study from South Africa, which compared the enterovirus genotypes in a small group of patients with severe and mild respiratory illness to asymptomatic individuals (94). Other enteroviruses, like EV-D68, have putatively been associated with respiratory illness and neurological symptoms (95). Enteroviruses are also the most common cause of aseptic viral meningitis, an inflammation of the tissue that covers the brain and the spinal cord. Although aseptic meningitis is mostly self-limiting, young children and immune-compromised individuals are more likely to have severe illness (1). One enterovirus genotype, E30, causes large outbreaks of aseptic meningitis, which is of public health concern (2,3). Therefore, it is important to identify and describe the genotypes causing illness or emerging genotypes. As recombination in enteroviruses is well documented there is an increasing need to produce full-length viral genome sequences (Figure 5.1). The ORF are flanked on either side by UTR's and the P1 genomic region encodes four capsid or structural proteins (VP1, VP2, VP3 and VP4) while the P2 and P3 genomic regions encode seven non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C and 3D).

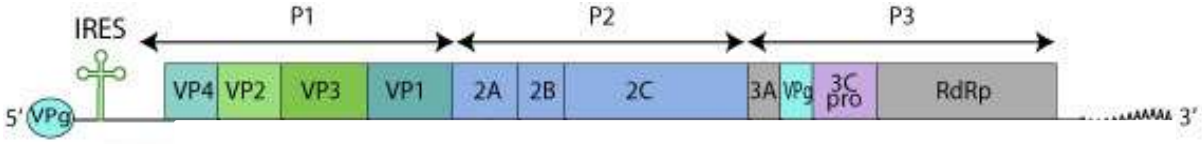


Figure 5.1: Genome structure of echovirus 30 (permission from ViralZone) (VPg – Viral protein genome linked; IRES – internal ribosomal entry site; RdRp – RNA dependent RNA polymerase)

Only nineteen E30 full length genome sequences are currently available in the public domain, mainly isolated from individuals with aseptic meningitis and only two isolated from individuals with respiratory illness in Thailand in 2010 (96). No full length E30 genome sequences are available from sub-Saharan Africa, or any countries with high HIV disease burden. E30 infections in HIV-infected individuals could lead to severe disease. Genome sequencing is important to identify evolutionary changes across the genome including potential inter-typic and intra-typic recombination events which might lead to emergence of contemporary strains with altered immune phenotypes or re-emergence of historic strains which may cause disease outbreaks (63,97).

Enrichment of samples for genome sequencing using next-generation sequencing can include RNA reverse transcription and cDNA amplification with random hexamers or alternatively with sequence independent single primer amplification (SISPA) primers (98,99). In some cases, the clinical samples can be filtered, treated with DNAse or subjected to a depletion of rRNA while in other cases combinations of these treatments can be used. The SISPA methodology for generation of whole genome sequences has not been described for enteroviruses before.

The aim of this study was to characterise E30 strains from respiratory illness cases across the complete viral genomes in South Africa from 2012 to 2014 using a SISPA-based next generation sequencing approach.

5.2 Materials and Methods

5.2.1 Study design and population

Study samples were obtained from participants enrolled in the SARI and ILI surveillance programmes as well as asymptomatic individuals from June 2012 through July 2014 (described in section 2.1).

5.2.2 Sample selection and detection of enterovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2.1. Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay which targets the conserved 5` untranslated region of enteroviruses (14) (methods section 2.2-2.3).

5.2.3 Amplification and sequencing of enteroviruses

The 5' proximal part of the VP1 region was amplified and sequenced, amino acid position 67 to position 133 relative to the start of the VP1 reading frame of the Fermon strain of enterovirus D68 (AY426531) (Section 2.4.1-2.4.3)(100).

5.2.4 Enterovirus isolation and genome sequencing

E30 positive samples were inoculated into human epithelial type 2 (HEp-2) cells for virus isolation and passage of cultures as described in section 2.4.4.

Rapid amplification of cDNA ends followed by sequence independent single primer amplification (RACE-SISPA) of E30 genomes from propagated virus were done with and without enrichment for viral RNA (section 2.4.5-2.4.6). Nineteen

reference E30 genome sequences downloaded are annotated with a GenBank accession number followed by place and year of isolation. GARD (https://www.datamonkey.org/) examined 10664 models, at a rate of 0.49 models per second to identify recombination breakpoints between the South African strains and other E30 genogroups (n=24). The alignment contained 2517 potential breakpoints, translating into a search space of 3.96364736471951e+22 models. Comparing the Akaike information criterion with a correction for small sample sizes (AICc) score of the best fitting GARD model, suggests that because the multiple tree model can be preferred over the single tree model by an evidence ratio of 100 or greater, at least of one of the breakpoints reflects a true topological incongruence. Usually the best model is chosen to be the model with the smallest AIC while the delta AIC calculates the difference between the AIC of each model (76).

5.3 Results

5.3.1 Enterovirus isolation

Of the genotyped samples, 9.9% (10/101) were previously identified as E30 by partial sequencing (Chapter 4). Only 8 of the 10 E30 positive samples had residual volume left for virus isolation and 5 (62.5%) were successfully propagated. E30 cultures presented with classic enterovirus cytopathic effect in Hep-2 cell monolayers (characterised by rounded, retractile cells detaching from the substrate surface with cytoplasmic stranding). Infection was confirmed by RT-PCR for enterovirus as described previously (100). All cultured samples were positive with Ct values less than 20.

5.3.2 Characteristics of the echovirus 30 positive cases

Of all 10 patients positive for E30, 20% (2/10) were enrolled with ILI and 80% (8/10) were enrolled with SARI. The majority (80%, 8/10) were paediatric patients less than 2 year of age.

All five successful E30 isolates were from children <2 years of age [mean 8.0 months (range, 1.4-18.4 months)]. Four of the five E30 isolates were derived from hospitalised patients who met the SARI case definition and were enrolled in 2013 at two surveillance sites; Edendale (25%, 1/4) and Klerksdorp (75%, 3/4). The fifth isolate was from a patient who met the ILI case definition and also enrolled in 2013 from an outpatient clinic in Klerksdorp. All five patients were HIV uninfected, but one hospitalised patient was HIV-exposed. The median duration of illness was 1 day before admission. The median duration of hospital stay was 4 days (range 3-37). Fever or history of fever was reported for all patients. None of the patients had any underlying illnesses. All patients were discharged and had an uneventful recovery (100).

5.3.3 Genome sequencing

Total nucleic acid samples enriched for RNA only before double stranded cDNA preparation

Only 1 (sample 1000411) of the TNA extracts from 5 E30 cultures had sufficient template concentration and acceptable quality (smear between 300-100 bp) for NGS following enrichment for viral RNA (Figure 5.2).

Non-enriched TNA samples for double stranded cDNA preparation

The non-enriched TNA samples' double stranded cDNA concentrations were as expected significantly higher ($p<0.001$) than those obtained for the RNA-enriched

samples (Table 5.1). All five samples were processed for NGS library preparation sequencing and reads obtained ranged from 2353047 to 3510297 (Table 5.1).

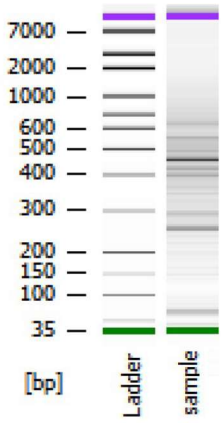


Figure 5.2: Example of electrophoresis of an echovirus 30 positive sample run on the Bioanalyzer, indicating smear between 300-1000bp

Table 5.1: Concentration of double stranded cDNA generated using RACE-SISPA comparing enriched and non-enriched viral samples as well as total reads generated with MiSeq

Sample ID	*TNA enriched for viral RNA [ng/µL]	**Non-enriched TNA [ng/µL]	Number of reads generated
502801	1.39	41.00	2481174
700714	1.55	33.60	2419220
1000411	3.73	110.30	2580378
700895	2.24	37.40	3510297
700919	1.61	20.00	2353047

*template for cDNA consists of RNA only; **template for cDNA consists of RNA and DNA

5.3.4 Complete genome analysis of echovirus 30 strains

The complete whole genomes were recovered by using reference mapping with the E30 prototype, Bastianni strain (AF311938) (Figure 5.3) after pairing and trimming of the total reads obtained for each sample as indicated in Table 5.1. The complete genome sequence consisting of 7453 nucleotides, of which 6582 nucleotides in an open reading frame encoding a polypeptide of 2194 amino acids

were recovered from the 5 E30 strains from South Africa. The coding sequences were flanked by a non-coding 5' UTR of 749 nucleotides and a non-coding 3' UTR of 99 nucleotides followed by a poly (A) tail composed of a long sequence of adenine nucleotides. Alignment of the full-length genome sequences from E30 isolates with the genome of the E30 Bastianni strain showed that they all had the same genomic organization and co-linear order of genomic regions.

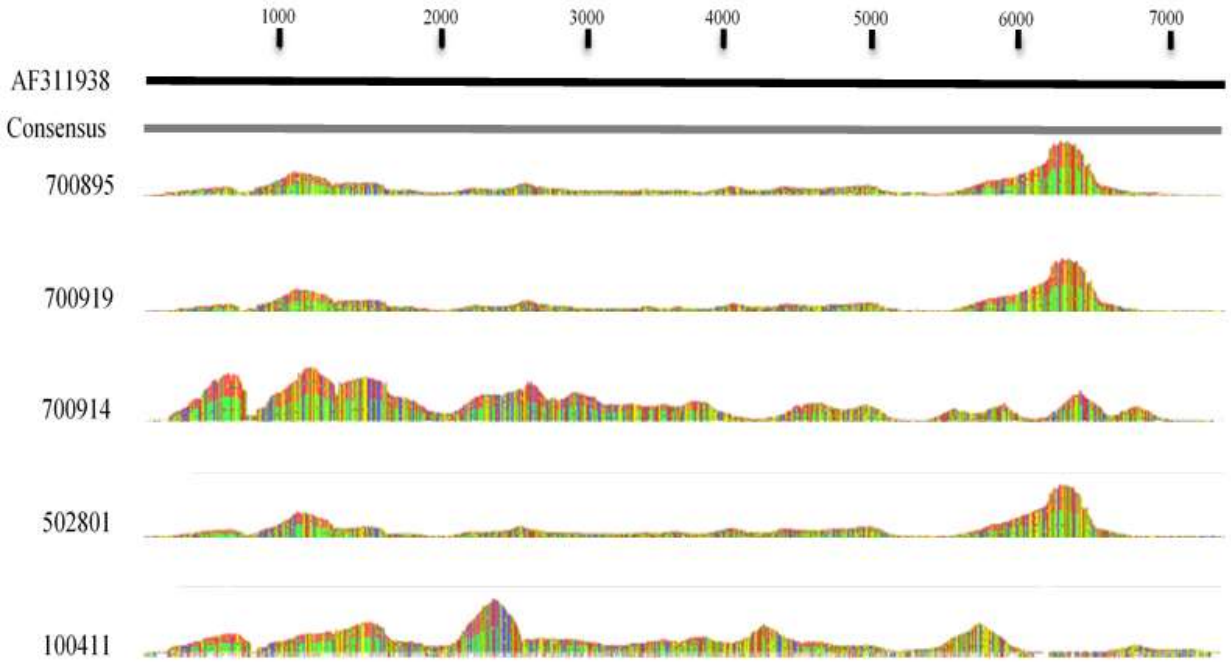


Figure 5.3: Schematic representation of genome coverage after reference mapping against the E30 prototype, Bastianni strain (AF311938)

Compared to the 19 E30 reference strains, the nucleotide sequences of the non-structural regions (P2 and P3) were more variable (74.4%-82.00% and 70.8%-85.1%, respectively), while the sequences of structural regions (P1) were more conserved (79.2%-85.4%) (Figure 5.4). This translated into amino acid changes and non-structural regions (P2 and P3) were more variable (76.5%-83.30% and 73.2%-86.2%, respectively), compared to the structural regions (P1) which were

more conserved (81.1%-86.4%). The complete genomes identified in this study have been deposited in GenBank (accession numbers MK376268-MK376272).

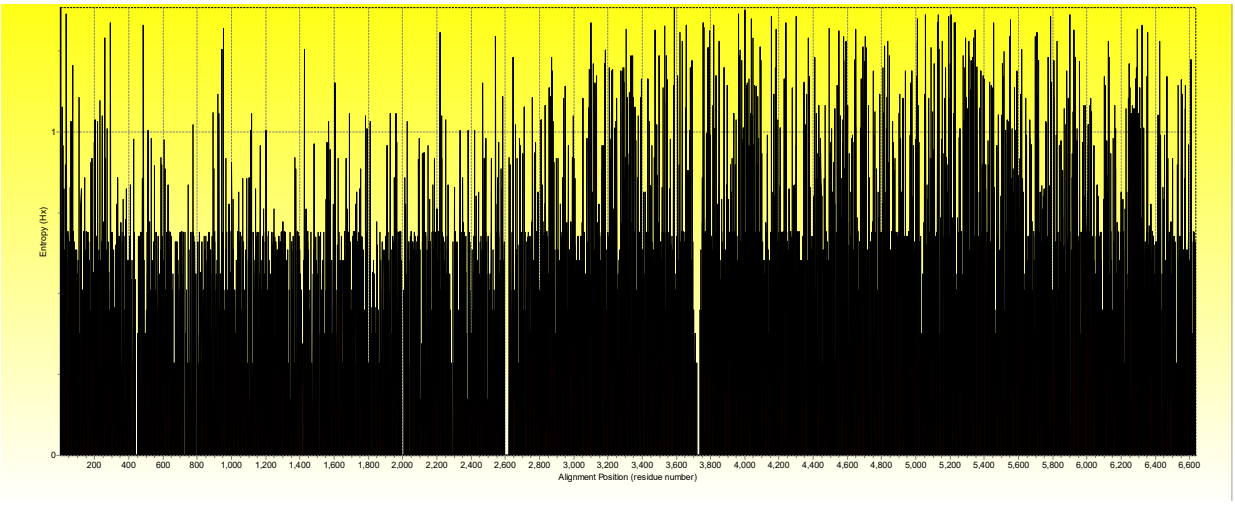


Figure 5.4: Entropy plot indicating the difference of the South African echovirus 30 strains across the polyprotein

5.3.5 Phylogenetic analysis

5.3.5.1 Phylogenetic analysis of complete polyprotein

Phylogenetic analysis of the polyprotein nucleotide sequences (6580 bp) displays a bi-phyletic clade supported by bootstraps of 72% and 94% respectively, for the South African E30 strains from 2013 (Figure 5.5). The South African viruses clustered with 94% bootstrap support with strains from the Netherlands, China and USA; where all South African strains are in a bi-phyletic sub-cluster supported by 100% bootstrap value. The complete genomes of South African E30 strains were also analysed with the online genotyping tool (https://www.rivm.nl/mpf/typingtool/enterovirus/) which could not assign a genotype.

Phylogenetic analysis were also done to assess if phylogenetic clustering observed for the complete polyprotein were maintained in the sub-genomic regions.

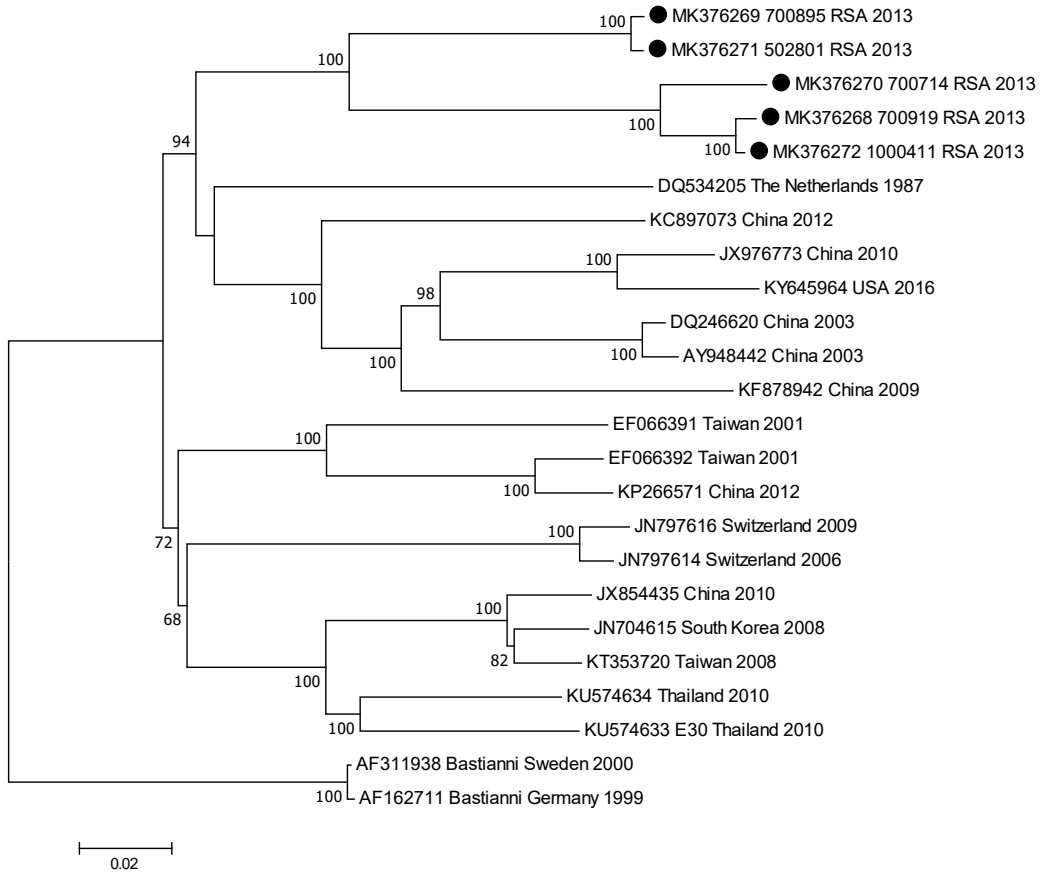


Figure 5.5: Genomic (6580 bp) relationship of 5 South African echovirus 30 (E30) strains compared to 19 other E30 strains available in GenBank database. Genome sequences were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstraps replicated). The scale bar represents the genetic distance. Bootstrap values >65% are shown. ● South African strains from this study.

5.3.5.2 Phylogenetic analysis of sub-genomic P1 region

The phylogenetic tree representing the P1 region (2580 bp) is indicated in Figure 5.6. The E30 strains from this study still clusters together with 100% bootstrap support.

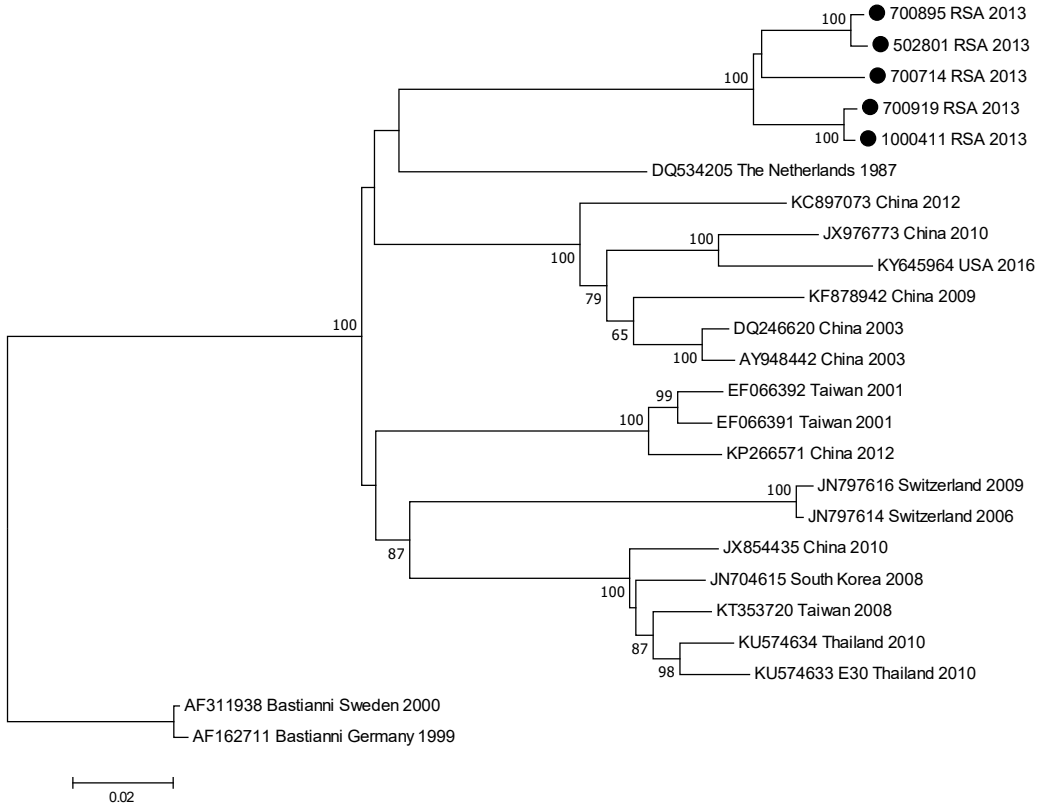


Figure 5.6: Phylogenetic relationships based on the complete P1 sub genomic region (2580 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

5.3.5.3 Phylogenetic analysis of sub-genomic P2 and P3 regions

The phylogenetic tree representing the P2 region is indicated in Figure 5.7. The South African strains did not form a single cluster, and two of the strains (MK376269 and MK376271) cluster with 70% bootstrap support with Asian strains from 2008 and 2010; including two strains (KU574633 and KU574634) from respiratory cases identified in Thailand in 2010. The other three South African strains form a separate sub cluster with strains from the Netherlands, China and USA.

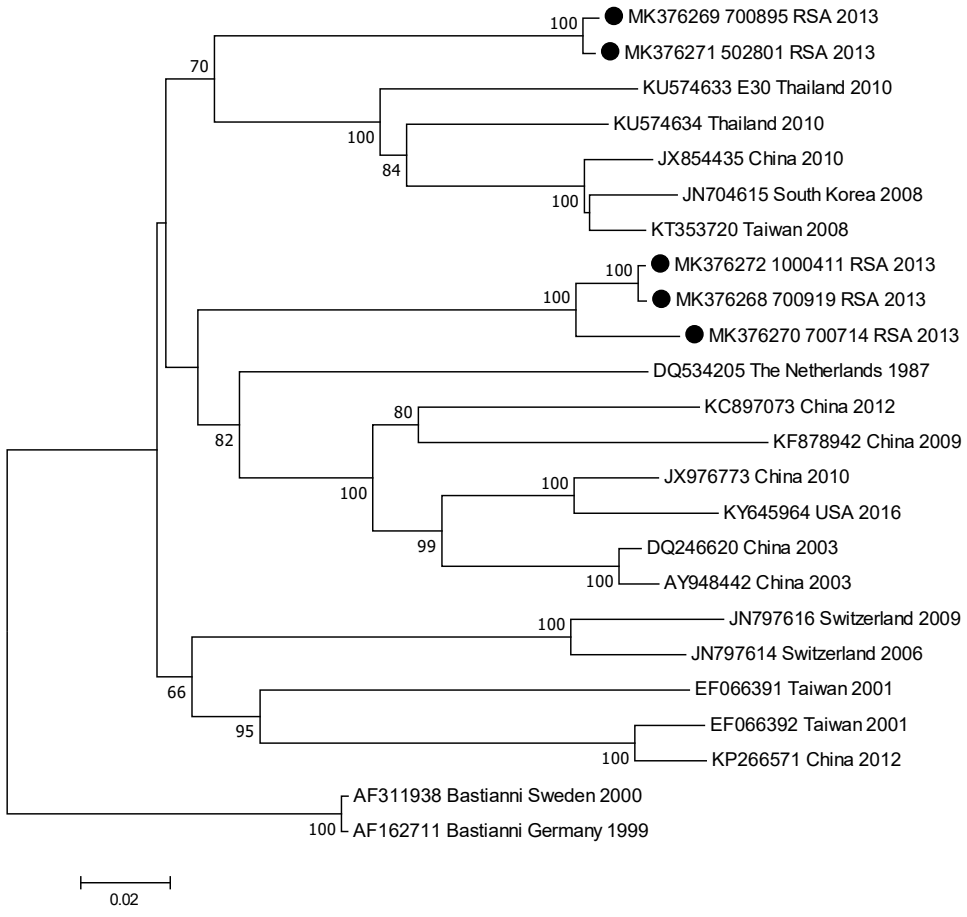


Figure 5.7: Phylogenetic relationships based on the complete P2 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

The phylogenetic tree representing the P3 region (2268 bp) is indicated in figure 5.8 and display similar clustering patterns for the South African E30 strains as was observed for the P2 region and the 2 subclusters are supported by 88% (MK376269 and MK376271) and 100% (MK376272, MK376268 and MK376270) bootstrap values respectively.

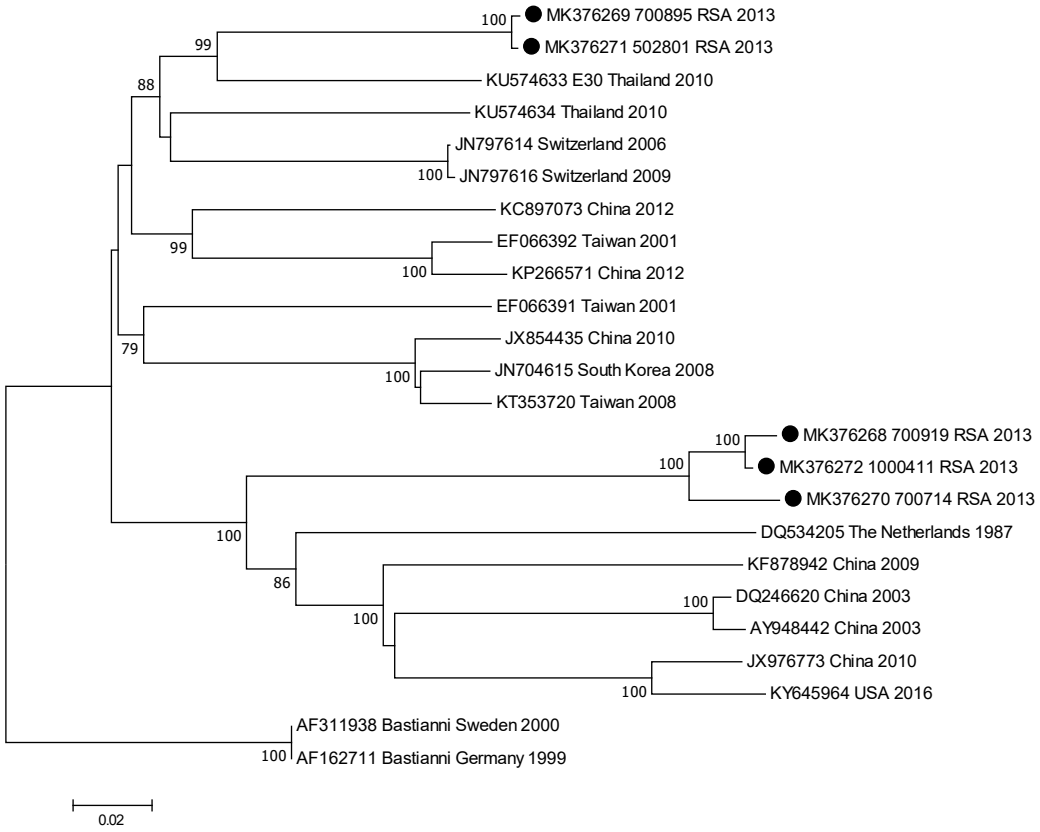


Figure 5.8: Phylogenetic relationships based on the complete P3 sub genomic region (1758 bp) of echovirus 30 (E30) strains. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study.

5.3.5.4 VP1 sub-genomic analysis

The phylogenetic tree representing the complete VP1 region (876 bp) is indicated in Figure 5.9. The five South African E30 strains still clustered separately (with 100% bootstrap support) to other E30 genogroups in a previously proposed genotype k.

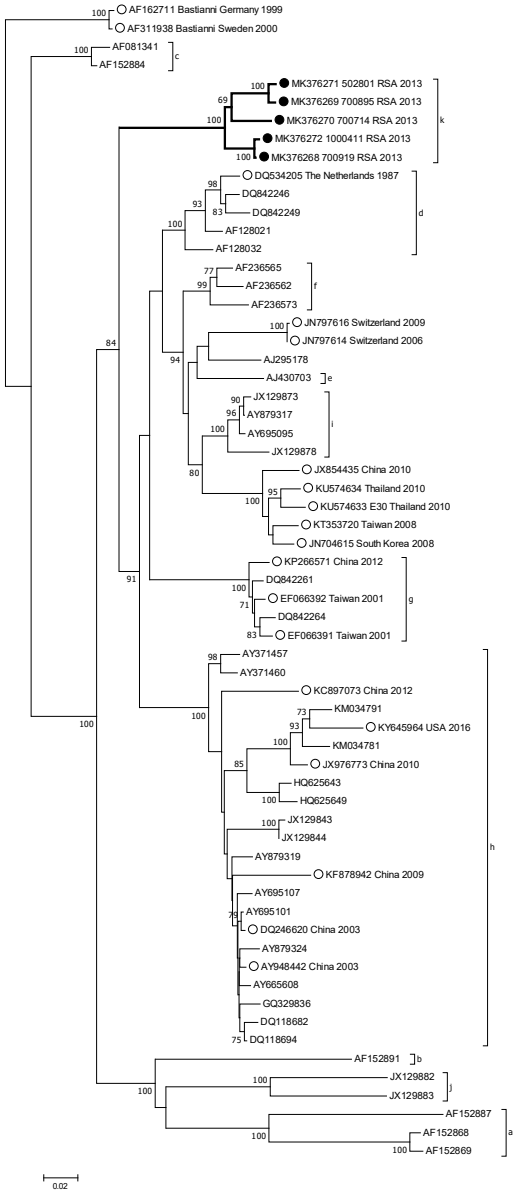


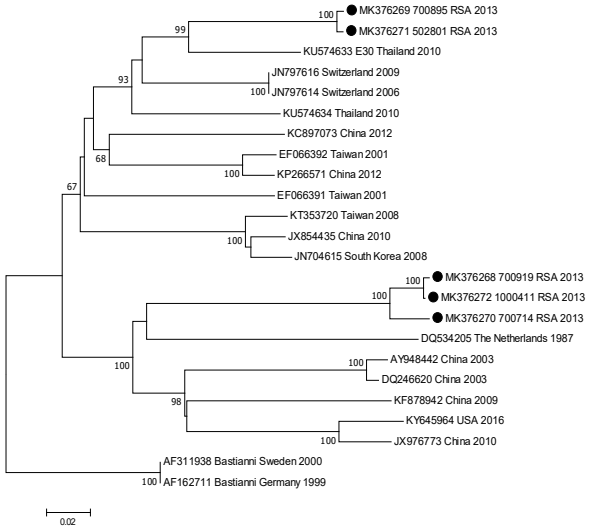
Figure 5.9: Phylogenetic relationships based on the complete VP1 genomic region (876 bp) of echovirus 30 (E30) strains. Five South African and 58 other E30 strains were analysed by nucleotide sequence alignment using the neighbor joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis. Bold line indicates the proposed genotype k.

5.3.6 Investigation of E30 genomes for recombination

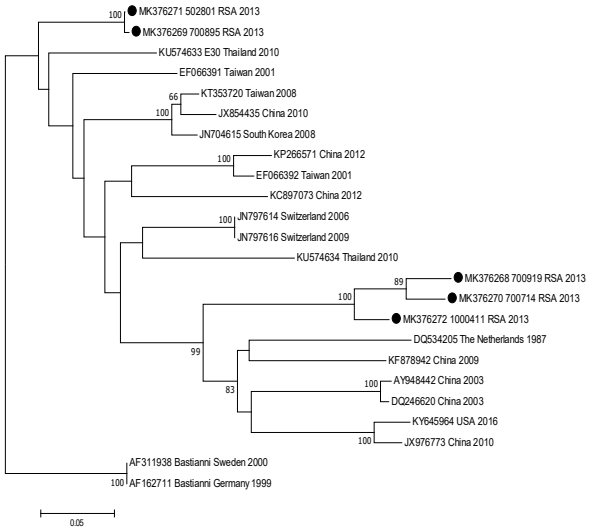
GARD identified at least 8 putative breakpoints of which breakpoints at nucleotide positions 2886, 4347, 5046 and 5364 were statistically significant (p= 0.0014, two-tailed; Figure 5.10). The phylogenetic trees for the various partitions at breakpoints confirms two distinct clusters in genomic regions outside of P1 and VP1. One cluster included two South African strains, MK376269 and MK376271 together with 2 European strains from 2006 and 2009 and 9 Asian strains ranging from 2001-2010. The other cluster consisted of three South African strains, MK376268, MK376272 and MK376270 together with 1 European strain from the 1980s, 4 Asian strains ranging from 2003-2010 and one recent USA strain (Figure 5.11). The clustering pattern of the South African strains did not change between the 4 partitions created by the significant breakpoints identified and confirms the distinct evolutionary history of the clusters. Further inter-typic recombination analysis should be done to identify the most recent common ancestor for the 2 clades in addition to generating more E30 genomes from various geographic regions and anatomical sites. Genomes were submitted to GenBank under accession numbers MK376268 to MK376272.

Topological incongruence report			
Breakpoints	LHS p-value	RHS p-value	Significance
2569	0.3654	0.0154	
2886	0.0014	0.0014	●●●●●●●●
3421	0.112	0.0014	
3768	0.8848	0.0014	
4347	0.0014	0.0014	●●●●●●●●
5046	0.0014	0.0014	●●●●●●●●
5364	0.0014	0.0014	●●●●●●●●

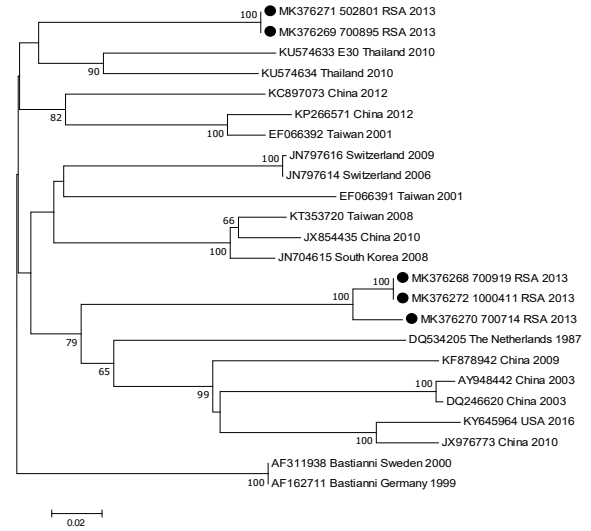
Figure 5.10: Topological incongruence report. LHS – left-hand side; RHS – right-hand side



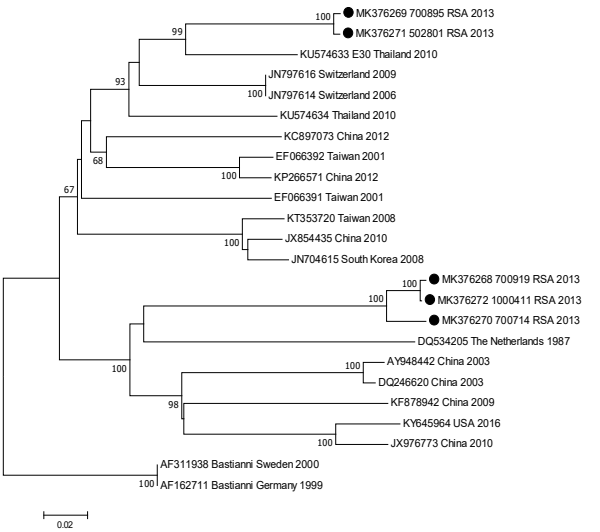
Nucleotide position 2886-4346



Nucleotide position 5046-5364



Nucleotide position 4347-5045



Nucleotide position 5364-6582

Figure 5.11: Phylogenetic relationships based on 4 partitions generated using the significant breakpoints across the echovirus 30 (E30) genome. Five South African and 19 other E30 strains were analysed by nucleotide sequence alignment using the neighbour-joining algorithm implemented in the MEGA 7.0.26 program. Numbers at the nodes indicate bootstrap support for that node (percentage of 1000 bootstrap replicated). The scale bar represents the genetic distance. Only high bootstrap values (>65%) are shown. ● South African strains from this study, E30 strains used for genomic and sub genomic analysis.

5.4 Discussion

This is the first study to describe the use of SISPA-PCR for whole genome sequencing of enteroviruses, specifically E30 viruses. It is known that the use of SISPA might decrease the possibility of specific primer bias and provides the opportunity to detect emerging clades or viruses with mutations in the primer binding sites (57). In this study, enrichment for viral RNA (DNase treatment) decreased the double stranded cDNA concentration obtained during RACE-SISPA between 10-30 times. From these results, the sample with the highest yield did not produce the most reads. It suggests that if the input template concentrations are too low, it affects the efficiency of the NGS library preparation step. Addition of carrier RNA to the low concentration enriched RNA templates may have resulted in successful NGS, but we did not explore this. This finding is similar to that of Hall *et al*., (101) which highlights the need for more data to validate the effect of enrichment methods on different sample types, amplification, bioinformatics approaches and sequencing platforms.

The complete polyprotein nucleotide sequences of E30 strains from respiratory illness cases in South Africa, display a single biphyletic clade on genomic analysis with reference strains from mainly meningitis cases. The South African viruses clustered closely (94% bootstrap support) with strains from the Netherlands, China and USA - this clustering was sustained when analysis was performed on the sub-genomic P1region, which encodes the structural/capsid proteins of the E30 virus. These structural proteins are the primary antigens that induce protective immunity and are involved in receptor binding and therefore plays a role in immunity and infectivity (2). This indicates that the structural proteins, forming the outer capsid of the virus, were genetically distant from strains isolated from patients with aseptic meningitis.
Phylogenetic analysis was also performed on the P2 and P3 sub-genomic region – the non-structural proteins responsible for viral protease, shutoff of host protein synthesis, increased membrane permeability, inhibition of cellular secretory pathways, vesicle formation and nucleotide triphosphatase activity. For the non-structural sub-genomic P2 and P3 regions, South African strains did not cluster

together but form 2 distinct subclusters; two South African strains cluster closely with Asian strains from 2008 and 2010 of which two were isolated from respiratory cases from Thailand in 2010.

Complete full genome phylogenetic analysis supports the original designation of genotype k for the South African strains which was based on the partial VP1 region (Chapter 4); confirmed in this study with 100% bootstrap support during complete VP1 analysis. However further subgenomic analysis based on significant recombination breakpoints show distinct evolutionary histories for the five E30 strains which forms 2 distinct South African clades as also described in sub-genomic analysis of P2 and P3 regions.

Molecular epidemiology studies of E30 meningitis outbreaks have shown that point mutations give rise to substantial genetic diversity in the capsid genes, specifically the VP1 polypeptide (2). The unique clustering of the South African E30 strains from respiratory specimens could be due to specific mutations in the capsid genes, although this was not investigated as part of this study. Complete VP1 and full genome phylogenetic analysis support the designation of a new genotype for South African E30 strains and specifically being designated genotype k as proposed previously. The fact that there are so little whole genome E30 sequences available from respiratory samples, could also be due to the fact that not a lot of research has been done on E30 viruses as a cause of respiratory illness.

Similar clustering was observed in the genomic polyprotein and P1 sub-genomic regions, but clustering differed significantly between other sub-genomic regions, similar results have been reported in a number of other studies (102). This could be due to different evolutionary paths of specific genes, an example of a gene symbiosis system, where genes, not species, are the elementary units of evolution. Distinct grouping was confirmed that could be a result of the diversity of different subgenomic regions in E30 driven by different evolutionary paths. This study only looked at intra-genotypic recombination and not at inter-genotypic recombination

that are generally observed in the non-structural regions (97). It was previously shown that recombination increases with geographical separation and time with recombination half-live estimated at 3.1 years for E30 (2,103). It is recommended that more full genome E30 positive samples from South Africa be sequenced to complement the five full E30 genomes submitted. As more complete genome sequence data from circulating E30 become available, the role of recombination on the evolutionary pathways of E30 may become apparent.

From this study, five E30 whole genome sequences were submitted to the public domain, and this contribution will increase the pool of whole genome sequences from sub-Saharan Africa. The majority of E30 genome data in the public domain are from patients with aseptic meningitis identified in Asia, USA and Europe ranging from 1987 to 2012, which may explain the genetic distance of these viruses compared to viruses from respiratory samples in this study.

Ongoing molecular surveillance is important to help identify emerging virus strains of clinical significance and to better understand circulation patterns of these viruses.

Chapter 6

The Fraction of Rhinovirus Detections Attributable To Mild and Severe Respiratory Illness in a Setting of High Human Immunodeficiency Virus Prevalence, South Africa, 2013–2015

Reference:

Hellferscee O, Treurnicht FK, Walaza S, du Plessis M, von Gottberg A, Moyes J, Dawood H, Variava E, Pretorius M, Venter M, Cohen C, Tempia S. The fraction of rhinovirus detections attributable to mild and severe respiratory illness in a setting of high human immunodeficiency virus prevalence, South Africa, 2013–2015. Journal of Infectious Diseases 2019 May 5;219(11):1697-1704. Doi:10.1093/infdis/jiy725

6.1 Introduction

In the family *Picornaviridae*, the genus Enterovirus includes rhinovirus and enterovirus species (104). Rhinoviruses are amongst the most commonly detected viruses in patients with upper respiratory tract infection and they are considered to be associated, in most cases, with mild, self-limiting cold-like illnesses (41,42). However, in studies conducted in South Africa, rhinoviruses were detected in 25%-34% of upper respiratory tract specimens from patients hospitalised with SARI (21,105).

Recent advances in the development of sensitive PCR assays have considerably expanded the ability of laboratories to detect pathogens, especially viral agents. Nonetheless, establishing a clinical association between pathogen detection and illness remains challenging. Improving the understanding of the fraction of rhinoviruses detection associated with mild or severe respiratory illness would provide insights for the interpretation of surveillance data, including the burden of illness associated with rhinoviruses infection.

This study aimed to estimate the fraction of rhinovirus detection attributable to illness (hereafter defined as RV-AF) among HIV-infected and HIV-uninfected South African patients of different age groups separately during 2013-2015. This was done among outpatients with ILI and inpatients with severe respiratory illness (SRI). In addition, the proportion of ILI and SRI cases attributable to rhinoviruses infection after adjusting for the RV-AF were estimated.

6.2 Materials and Methods

6.2.1 Study design and population

Study samples were obtained from participants enrolled in the SRI (SARI and SCRI combined) and ILI surveillance programmes as well as asymptomatic individuals from June 2013 through July 2015 (described in section 2.1). For this study, 2012 samples were excluded as control recruitment started late and we therefore extended the study period to 2015 instead of 2014.

6.2.2 Sample selection and detection of rhinovirus infection

Respiratory specimens were collected from all enrolled patients and controls as described in section 2.2-2.3.

Specimens were tested for the presence of 10 respiratory viruses (enterovirus, influenza A and B viruses; parainfluenza virus types 1, 2 and 3; respiratory syncytial virus; adenovirus; rhinovirus; human metapneumovirus) using a multiplex qRT-PCR assay (14) (methods section 2.4.1-2.4.2) during 2013-2014. In 2015, a combination of the commercial Fast Track Diagnostics influenza virus and respiratory syncytial virus assay (Fast Track Diagnostics, Luxembourg) and Allplex Respiratory Assay, panel 2 and 3 (Seegene, Seoul, Korea) was used. Human bocavirus, human coronaviruses (OC43, HK1, NL63, 229E) and para-influenza virus 4 were only tested in 2015 and were excluded from the analysis. The platform was changed to a commercial assay because the in-house multiplex PCR was no-longer available and detection of all 10 pathogens were validated (with samples containing high and low viral loads) before implementing the commercial assay.

6.2.3 Statistical analysis

As described in section 2.5.

6.3 Results

6.3.1 Study population

Over the study period, 10,576 individuals were enrolled, of whom 10,134 (95.8%) had known age and available rhinovirus and HIV results, and were thus included for further analysis. From the 10,134 study participants, 2120 (20.9%) were controls, 3654 (36.1%) had ILI; and 4360 (43.0%) had SRI. Of the 4360 patients with SRI, 2735 (62.7%) had SARI and 1625 (37.3%) had SCRI. Children aged <5 years accounted for 38.0% (806/2120) of controls, 35.5% (1297/3654) of patients with ILI and 37.4% (1630/4360) of patients with SRI [55.6% (1520/2735) of patients with SARI; and 6.7% (110/1625) of patients with SCRI].

Overall, the HIV prevalence was 41.2% (874/2120) among controls (owing to enrolment criteria for controls); 25.0% (915/3654) among patients with ILI and 50.1% (2186/4360) among patients with SRI [37.1% (1015/2735) among patients with SARI; and 72.1% (1171/1625) among patients with SCRI ($p<0.001$)]. Across syndromes, the HIV prevalence was lowest among infants aged <1 year [1.5% (7/464) among those with ILI; 9.3% (96/1032) among those with SRI; 8.7% (84/961) among those with SARI; and 16.9% (12/71) among those with SCRI], and it was highest among patients aged 25–44 years [59.3% (581/980) among those with ILI, 90.5% (1221/1349) among those with SRI; 89.3% (509/570) among those with SARI; and 91.4% (712/779) among those with SCRI] (Table 6.1).

Table 6.1: HIV prevalence among outpatients with influenza-like illness and inpatients with severe (acute or chronic) respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013-December 2015

Age, year	Influenza-like illness	Severe respiratory illness	Severe acute respiratory illness	Severe chronic respiratory illness
	% (n/N)	% (n/N)	% (n/N)	% (n/N)
<1	1.5 (7/464)	9.3 (96/1032)	8.7 (84/961)	16.9 (12/71)
1–4	2.0 (17/833)	13.7 (82/598)	12.7 (71/559)	28.2 (11/39)
5–24	15.7 (152/970)	60.4 (198/328)	54.3 (102/188)	68.6 (96/140)
25–44	59.3 (581/980)	90.5 (1221/1349)	89.3 (509/570)	91.4 (712/779)
45–64	43.9 (154/351)	64.9 (535/824)	65.6 (225/343)	64.5 (310/481)
≥65	7.1 (4/56)	23.6 (54/229)	21.1 (24/114)	26.1 (30/115)
Total	25.0 (915/3654)	50.1 (2186/4360)	37.1 (1015/2735)	72.1 (1171/1625)

6.3.2 Rhinovirus detection and attributable fraction

In all age groups during the study period, rhinoviruses were detected in 23.2% (2350/10134) of specimens from enrolled study participants; 17.4% (368/2120) among controls (Table 6.2), 26.8% (979/3654) among patients with ILI (Table 6.3), and 23.0% (1003/4360) among patients with SRI ($p<0.001$) (Table 6.4) [27.5% (751/2735) among those with SARI (Table 6.5); and 15.5% (252/1625) among those with SCRI (Table 6.6)].

The overall RV-AF was 54.6% (95% CI: 47.5-60.8%) among patients with ILI (Table 6.3, Figure 6.1A) and 51.4% (95% CI: 43.4-58.3%) among patients with SRI (Table 6.4, Figure 6.2A) [54.3% (95% CI: 46.3-61.1%) among those with SARI (Table 6.5); and 37.5% (95% CI: 21.5-50.2%) among those with SCRI (Table 6.6)]. The RV-AF was significant among both HIV-infected and HIV-uninfected individuals with ILI (HIV-infected: 59.9%, 95% CI: 45.8-70.3% vs HIV-uninfected: 53.6%, 95% CI: 44.7-61.1%) or SRI (HIV-infected: 39.8%, 95% CI: 22.3-53.3% vs HIV-uninfected: 55.3%, 95% CI: 45.6-63.2%) (Tables 6.3 and 6.4, Figure 6.1B-C, Figure 6.2B-C). The RV-AF was also significant among HIV-infected and HIV-uninfected individuals aged <5 and ≥5 years with ILI or SRI, with the exception of HIV-infected children aged <5 years with SRI (Tables 6.3 and 6.4). Non-significant RV-AF were observed in some of the other age groups- and HIV-stratified analyses.

Overall, the RV AF-adjusted prevalence was 14.6% among patients with ILI (Table 6.3, Figure 6.1A) and 11.8% among patients with SRI (Table 6.4, Figure 6.2A) [14.9% among those with SARI (Table 6.5); and 5.8% among those with SCRI (Table 6.6)].

Table 6.2. Asymptomatic individuals positive for rhinovirus infection, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Age, year	All no. positive/no. total (%)	HIV-infected no. positive/no. total (%)	HIV-uninfected no. positive/no. total (%)
<1	98/339 (28.9%)	7/31 (22.6%)	91/308 (29.5%)
1–4	121/467 (25.9%)	53/205 (25.9%)	68/262 (26.0%)
5–24	94/592 (15.9%)	43/280 (15.4%)	51/312 (16.3%)
25–44	27/310 (8.7%)	17/189 (9.0%)	10/121 (8.3%)
45–64	18/280 (6.4%)	12/141 (8.5%)	6/139 (4.3%)
≥65	10/132 (7.6%)	3/28 (10.7%)	7/104 (6.7%)
<5	219/806 (27.2%)	60/236 (25.4%)	159/570 (27.9%)
≥5	149/1314 (11.3%)	75/638 (11.8%)	74/676 (10.9%)
All	368/2120 (17.4%)	135/874 (15.4%)	233/1246 (18.7%)

Table 6.3. Rhinoviruses detection and rhinoviruses attributable fraction among outpatients with influenza-like illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Outpatients with influenza-like illness									
	All			HIV-infected			HIV-uninfected		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%
<1	**48.1 (28.6 to 62.3)**	181/464 (39.0)	18.8	96.5 (0 to 99.9)	3/7 (42.9)	41.4	**46.7 (26.4 to 61.5)**	178/457 (38.9)	18.2
1–4	**40.8 (18.9 to 56.8)**	269/833 (32.3)	13.2	54.3 (0 to 86.0)	6/17 (35.3)	19.2	**40.1 (17.0 to 56.8)**	263/816 (32.2)	12.9
5–24	**53.9 (38.8 to 65.2)**	248/970 (25.6)	13.8	**40.9 (0.9 to 64.8)**	34/152 (22.4)	9.2	**58.8 (41.9 to 70.8)**	214/818 (26.2)	15.4
25–44	**68.7 (52.2 to 79.5)**	204/980 (20.8)	14.3	**66.0 (41.8 to 80.1)**	119/581 (20.5)	13.5	**72.6 (46.0 to 86.2)**	85/399 (21.3)	15.5
45–64	**73.5 (54.2 to 84.6)**	68/351 (19.4)	14.3	**70.5 (41.0 to 85.2)**	36/154 (23.4)	16.5	**76.7 (44.0 to 90.3)**	32/197 (16.2)	12.4
≥65	61.7 (0 to 85.6)	9/56 (16.1)	9.9	84.4 (0 to 98.8)	1/4 (25)	21.1	59.9 (0 to 86.0)	8/52 (15.4)	9.2
<5	**44.6 (30.7 to 55.7)**	450/1297 (34.7)	15.5	**68.6 (10.6 to 88.98)**	9/24 (37.5)	25.7	**43.6 (29.1 to 55.2)**	441/1273 (34.6)	15.1
≥5	**62.9 (54.4 to 69.8)**	529/2357 (22.4)	14.1	**59.5 (44.5 to 70.49)**	190/891 (21.3)	12.7	**64.4 (53.0 to 73.0)**	339/1466 (23.1)	14.9
All	**54.6 (47.5 to 60.8)**	979/3654 (26.8)	14.6	**59.9 (45.8 to 70.3)**	199/915 (21.7)	13.1	**53.6 (44.7 to 61.1)**	780/2739 (28.5)	15.3

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in non-stratified models.

Table 6.4. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe respiratory illness									
All				**HIV-infected**			**HIV-uninfected**		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%
<1	**38.8 (17.6 to 54.5)**	332/1032 (32.2)	12.5	18.6 (0 to 68.8)	27/96 (28.1)	5.2	**41.5 (20.0 to 57.3)**	305/936 (32.6)	13.5
1–4	**60.2 (45.8 to 70.7)**	241/598 (40.3)	24.3	20.3 (0 to 58.0)	30/82 (36.6)	7.4	**68.4 (54.6 to 77.9)**	211/516 (40.9)	28
5–24	**55.2 (37.2 to 68.0)**	96/328 (29.3)	16.2	**45.5 (13.5 to 65.6)**	53/198 (26.8)	12.2	**64.0 (41.3 to 78.0)**	43/130 (33.1)	21.2
25–44	**47.4 (18.8 to 65.9)**	219/1349 (16.2)	7.7	**51.1 (18.1 to 70.8)**	204/1221 (16.7)	8.5	30.4 (0 to 69.9)	15/128 (11.7)	3.6
45–64	**43.8 (5.0 to 66.7)**	92/824 (11.2)	4.9	26.5 (0 to 61.2)	62/535 (11.6)	3.1	**60.4 (5.0 to 83.5)**	30/289 (10.4)	6.3
≥65	30.5 (0 to 68.0)	23/229 (10.0)	3.1	-23.7 (0 to 71.7)	4/54 (7.4)	0	42.7 (0 to 76.7)	19/175 (10.9)	4.7
<5	**50.3 (38.6 to 59.9)**	573/1630 (35.2)	15.7	21.2 (0 to 53.8)	57/178 (32.0)	6.8	**56.0 (44.3 to 65.3)**	516/1452 (35.5)	19.9
≥5	**51.3 (38.7 to 61.3)**	430/2730 (15.8)	8.1	**44.8 (25.6 to 59.1)**	323/2008 (16.1)	7.2	**57.9 (39.9 to 70.5)**	107/722 (14.8)	8.6
All	**51.4 (43.4 to 58.3)**	1003/4360 (23.0)	11.8	**39.8 (22.3 to 53.3)**	380/2186 (17.4)	6.9	**55.3 (45.6 to 63.2)**	623/2174 (28.7)	1.9

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

Table 6.5. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe acute respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe acute respiratory illness									
	Overall			**HIV-infected**			**HIV-uninfected**		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%	% (95% CI)	n/N (%)	%
<1	**41.1 (20.4 to 56.4)**	312/961 (32.5)	13.4	29.4 (0 to 73.3)	26/84 (31.0)	9.1	**43.2 (22.0 to 58.7)**	286/877 (32.6)	14.1
1–4	**60.5 (45.8 to 71.2)**	224/559 (40.1)	24.3	7.3 (0 to 53.1)	24/71 (33.8)	2.5	**69.2 (55.6 to 78.7)**	200/488 (41.0)	28.4
5–24	**61.8 (43.7 to 74.1)**	61/188 (32.4)	20.1	**43.6 (1.3 to 67.8)**	26/102 (25.5)	11.1	**73.7 (54.6 to 84.8)**	35/86 (40.7)	30
25–44	**55.8 (29.4 to 72.3)**	103/570 (18.1)	10.1	**59.6 (30.5 to 76.5)**	96/509 (18.9)	11.3	38.1 (0 to77.2)	7/61 (11.5)	3.9
45–64	**49.1 (8.5 to 71.7)**	40/343 (11.7)	5.7	14.3 (0 to 58.5)	24/225 (10.7)	1.5	**73.2 (30.3 to 89.7)**	16/118 (13.6)	9.9
≥65	27.7 (0 to 71.1)	11/114 (9.6)	2.6	0 (0 to 75.2)	0/24 (0)	0	50.0 (0 to 81.8)	11/90 (12.2)	6
<5	**51.6 (39.9 to 61.0)**	536/1520 (35.3)	18.2	16.4 (-45.3 to 51.9)	50/155 (32.3)	5.3	**57.2 (45.5 to 66.3)**	486/1365 (35.6)	20.4
≥5	**58.0 (45.7 to 67.6)**	215/1215 (17.7)	10.3	**45.4 (23.8 to 60.9)**	146/860 (17.0)	7.7	**68.8 (53.7 to 79.0)**	69/355 (19.4)	13.3
All	**54.3 (46.3 to 61.1)**	751/2735 (27.5)	14.9	**39.0 (19.3 to 54.0)**	196/1015 (19.3)	7.5	**59.5 (50.3 to 67.0)**	555/1720 (32.3)	19.2

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

Table 6.6. Rhinoviruses detection and rhinoviruses attributable fraction among patients hospitalised with severe chronic respiratory illness by HIV infection status, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015

Inpatients with severe chronic respiratory illness									
	Overall			HIV-infected			HIV-uninfected		
Age, year†	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate	Attributable fraction	Observed detection rate	Adjusted detection rate
	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%	(95% CI)	n/N (%)	%
<1	0 (0 to 45.0)	20/71 (28.2)	-0.5	0 (0 to 69.4)	1/12 (8.3)	-8.9	17.8 (0 to 57.9)	19/59 (32.2)	5.7
1–4	**54.9 (9.4 to 77.6)**	17/39 (43.6)	23.9	63.5 (0 to 90.1)	6/12 (50.0)	34.7	52.1 (0 to 79.3)	11/28 (39.3)	20.5
5–24	**39.6 (3.9 to 62.1)**	35/140 (25.0)	9.9	**46.6 (5.1 to 70.0)**	27/96 (28.1)	13.1	32.1 (0 to 70.6)	8/44 (18.2)	5.8
25–44	**41.6 (7.0 to 63.3)**	116/779 (14.9)	6.2	**43.7 (4.0 to 67.0)**	108/712 (15.2)	6.6	28.2 (0 to 73.3)	8/67 (11.9)	3.4
45–64	41.1 (0 to 66.6)	52/481 (10.8)	4.4	38.8 (0 to 69.4)	38/310 (12.3)	4.8	44.8 (0 to 78.8)	14/171 (8.2)	3.7
≥65	34.0 (0 to 72.5)	12/115 (10.4)	3.5	33.7 (0 to 85.2)	4/30 (13.3)	4.5	34.5 (0 to 77.1)	8/85 (9.4)	3.2
<5	26.8 (0 to 53.7)	37/110 (33.6)	9	26.4 (0 to 73.6)	7/24 (30.4)	8	34.4 (0 to 61.0)	30/87 (34.5)	11.9
≥5	**39.9 (21.5 to 54.0)**	215/1515 (14.2)	5.7	**44.6 (22.7 to 60.3)**	177/1148 (15.4)	6.9	33.7 (0 to 58.6)	38/367 (10.4)	3.5
All	**37.5 (21.5 to 50.2)**	252/1625 (15.5)	5.8	**42.9 (21.6 to 58.3)**	184/1172 (15.7)	6.7	**33.6 (6.34 to 53.0)**	68/454 (15.0)	5

Bold values indicates significant values †The overall rhinovirus AF was obtained from models adjusted for any underlying medical conditions and for other respiratory viruses investigated in this study. Age group and HIV infection were also included as predictors in nonstratified models.

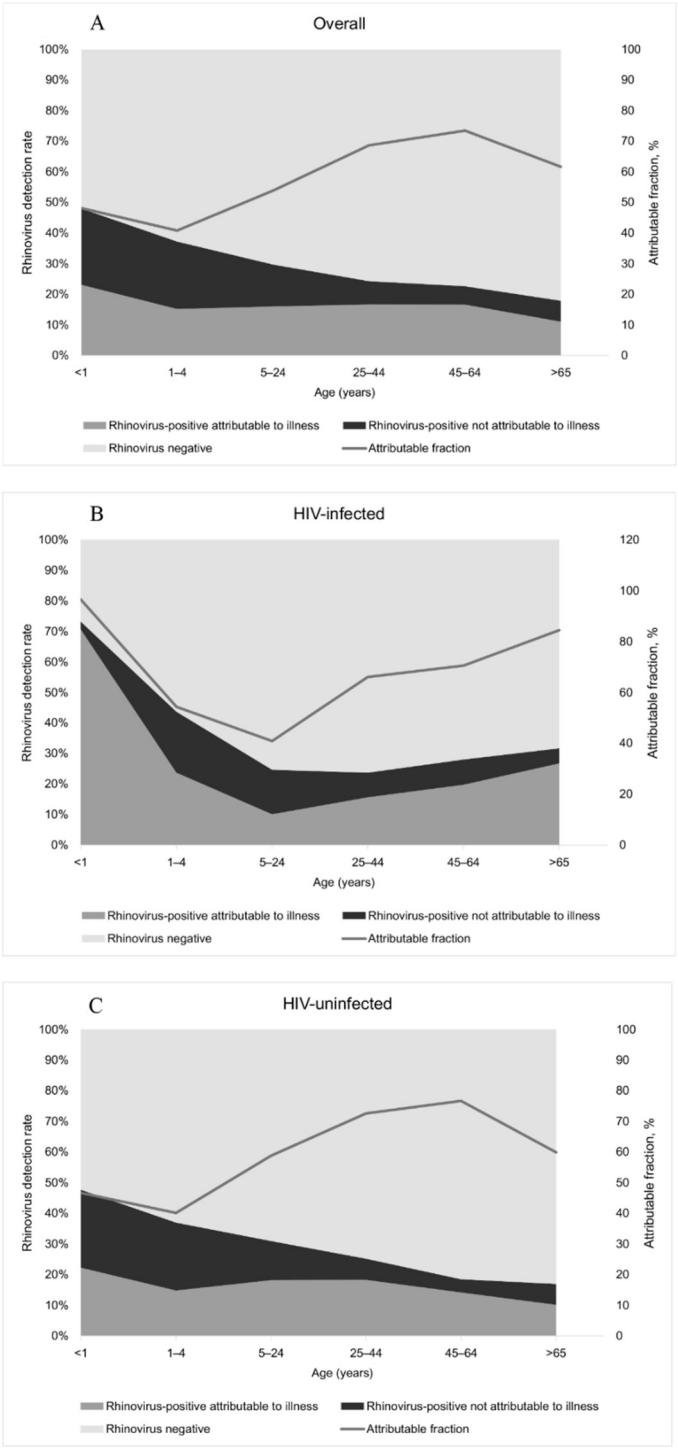


Figure 6.1: Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among outpatients with influenza-like illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected patients.

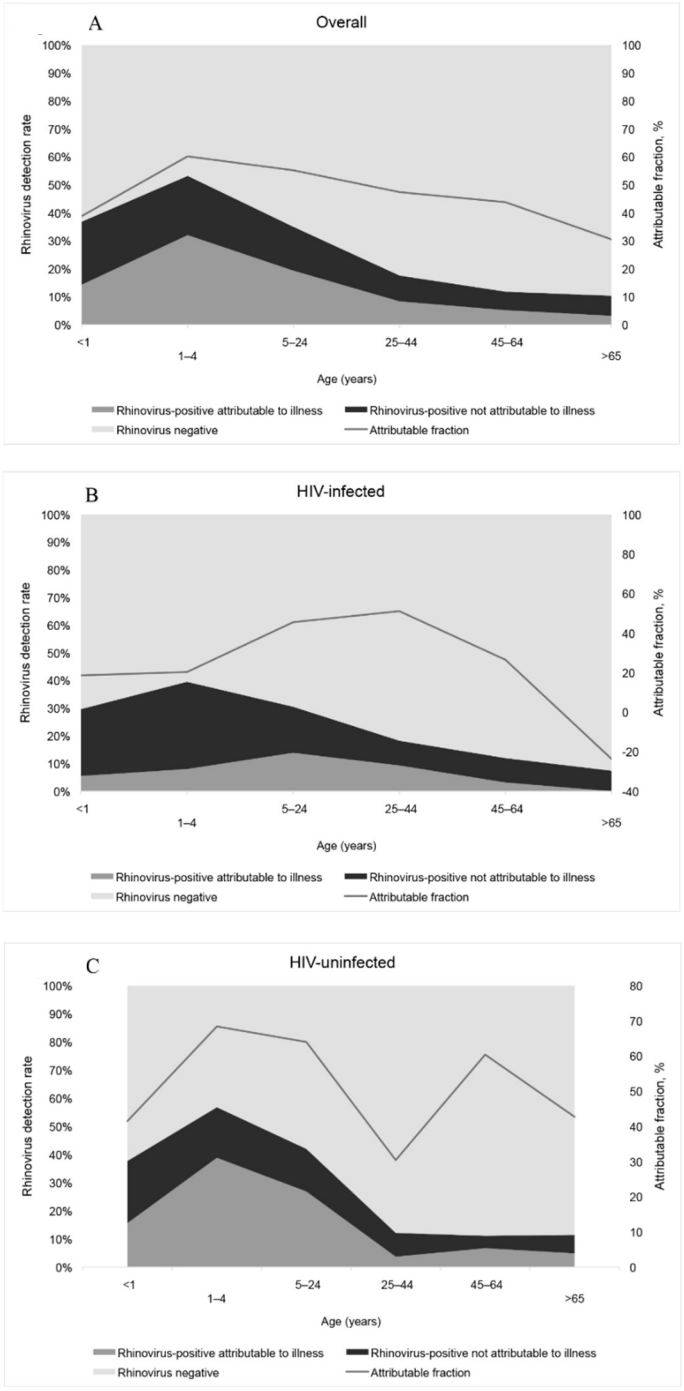


Figure 6.2: Estimated rhinovirus attributable fraction and proportion of rhinovirus-positive cases attributable and not attributable to illness among patients hospitalised with severe respiratory illness, Klerksdorp and Pietermaritzburg, South Africa, January 2013–December 2015. (A) Overall. (B) HIV-infected patients. (C) HIV-uninfected

6.4 Discussion

This study assessed the association between rhinovirus detection and mild or severe respiratory illness, among HIV-infected and HIV-uninfected patients of different age groups separately. Overall, rhinovirus was significantly associated with illness across syndromes and HIV status; however, this association was not significant in some age groups. The estimated detection rate attributable to illness reported in this study reflects a more accurate measure of the proportion of rhinovirus detections causing respiratory illness in both children and adults in South Africa, contrary to reporting viral detection rates alone. This study confirmed an overall significant RV-AF among patients with mild or severe respiratory illness as previously described in South Africa (21).

The overall RV-AF (ILI: 54.6%; SARI: 54.3%) found in this study were similar to those reported in a previous study conducted in South Africa (ILI: 52.0%; SARI: 46.9%), but was lower compared to those of other important respiratory pathogens such as influenza viruses (ILI: 93.3%; SARI: 86.3%) and respiratory syncytial virus (ILI: 63.1%; SARI: 83.7%) (21). Although this study found a significant RV-AF among HIV-infected and HIV-uninfected individuals across syndromes, only approximately half of the illness among RV-positive patients would be actually attributable to rhinovirus infection.

It was hypothesized that the elevated HIV infection prevalence in the South African setting may have been responsible for the significant attributable fraction found in a previous South African study (21); however, this study observed that rhinovirus detection was significantly associated with both mild and severe respiratory illness also among HIV-uninfected individuals. Studies from Kenya, Mali, Thailand and USA did not find an association of rhinovirus infection with illness (47–52,106); however, these studies had some limitations, including tests that combined rhinovirus and enterovirus detection and small samples sizes. The significant RV-AF found in this study among both HIV-infected and HIV-uninfected individuals may be due to increased power associated with the large sample size of this study compared to the above-mentioned studies. After

adjusting for the RV-AF the proportion of illness attributable to rhinovirus (ILI: 11.8%; SARI: 14.9%) was similar to or higher than those of influenza viruses (ILI: 14.2%; SARI: 4.9%) or respiratory syncytial virus (ILI: 5.9%; SARI: 16.7%) (21,107). This suggests that rhinovirus, whereas not always associated with illness when detected, may still be responsible for a substantial number of mild and severe respiratory illness cases.

This study has limitations that warrant discussion. Firstly, rhinovirus nucleic acids could be present during the pre- or post-syndromic phase of the infection and can be detected using PCR. In addition, controls after enrolment were not followed-up, and the development of undetected mild or severe illness for some individual cannot be excluded. The inclusion in the control group of individuals with rhinovirus -associated symptomatic illness during the pre- or post-syndromic phase would have resulted in an underestimation of the RV-AF. Secondly, the role of bacterial co-infections was not investigated in this analysis. Third, the small sample size in some of the age group- and HIV-stratified analyses may have hindered the ability to reach statistical significance in some of the groups. In addition, not having complete CD4+ count data, hindered the ability to assess the RV-AF among severely vs. mildly or not immunosuppressed HIV-infected individuals. Fourth, due to low numbers this study could not assess the RV-AF among patients with very severe illness, such as ICU admission or death. Last, no lower respiratory tract specimens were collected, hindering the ability to assess a differential rhinovirus positivity proportion between lower and upper respiratory tract specimens.

In conclusion, rhinoviruses were commonly identified among asymptomatic individuals but may still be associated with a substantial proportion of mild and severe respiratory illness across age groups and among HIV-infected and HIV-uninfected individuals. The proportion of illness attributable to rhinovirus after adjusting for the RV-AF was similar to or higher than those of other important pathogens such as influenza viruses and respiratory syncytial virus. This has

potential implications on the prioritization of pathogens for vaccines and/or therapeutics development

Chapter 7

Conclusions

Enteroviruses are known causes of respiratory illness, but the epidemiology and contribution has not been well described in South Africa. In this study, the epidemiology and genotypes associated with SARI in South African patients were determined. Enterovirus genotypes associated with SARI were compared to those from milder ILI and asymptomatic controls to determine if certain enterovirus genotypes were more prevalent in mild or severe patients with respiratory illness. In this study enteroviruses frequently contributed to SARI and ILI and that there was a substantial diversity in enterovirus genotypes circulating in South African patients hospitalised with SARI. EV-D68, an enterovirus genotype of public health importance, circulated at high frequencies among children less than 5 years of age but did not contribute to more severe disease and adverse outcome. Unfortunately, the surveillance system did not record symptoms of neurological disease or paralysis and therefore this study was unable to investigate a causative link between EV-D68 infection and neurological disease.

Furthermore, this study did not identify any association between disease severity and different enterovirus genotypes when comparing mild (ILI) and severe (SARI) cases to asymptomatic controls. However, some genotypes (E30, coxsackie virus B5, EV-D68) were more prevalent in symptomatic cases. This study also reported that E30 viruses which were the most prevalent in the respiratory illness cases formed a distinct genetic cluster upon phylogenetic analysis of the sub genomic VP1 region which I proposed to call genotype k. However, requirements for new virus clade designations require that genome sequencing is done.

Therefore, in this study the, SISPA-based NGS was employed to sequence the complete E30 genomes, predominantly from paediatric SARI cases. I successfully sequenced five E30 genomes which were submitted to GenBank; contributing to the pool of whole genome sequences available for E30, especially from sub-Saharan Africa. No evidence of recombination could be confirmed, although distinct grouping was observed that could be a result of the diversity of different subgenomic regions in E30.

Lastly we also wanted to determine the epidemiology of rhinovirus species from the genus Enterovirus and assess the proportion of rhinovirus positives that contribute to clinical illness in the context of a high HIV prevalence setting. Previous studies reported that that HIV infection caused increased severity and higher case fatality ratios, especially in children (6,7). We determined the attributable fraction of rhinovirus infections among HIV-infected and HIV-uninfected patients with mild and severe respiratory illness in South Africa during 2013 to 2015. Rhinoviruses are the most prevalent species from the genus Enterovirus identified among patients hospitalised with SARI, but they are also like enterovirus commonly identified in healthy individuals. For this reason, it was decided to determine the fraction of rhinovirus detection associated with mild or severe respiratory illness. An overall significant attributable fraction of rhinoviruses among patients with mild or severe respiratory illness was confirmed, which suggest that rhinoviruses, whereas not always associated with illness, may still be responsible for a substantial number of mild and severe respiratory illness cases. The proportion of illness attributable to rhinoviruses (ILI: 11.8%; SARI: 14.9%) were similar to or higher than other pathogens of public health importance such as influenza viruses (ILI: 14.2%; SARI: 4.9%) and respiratory syncytial virus (ILI: 5.9%; SARI: 16.7%) (21).

In order to determine if other factors such as viral load or host characteristics play a role in enterovirus-associated disease, further studies will need to be conducted to increase the knowledge regarding the epidemiology of enteroviruses.

Currently, there is no vaccine or antiviral therapy available for non-polio enteroviruses and by understanding the contribution of enterovirus genotypes to severe illness this will allow for informed decision making on potential candidate vaccines and development of therapeutic interventions.

To conclude, study findings show that (i) enterovirus genotypes detected in mild versus severe disease do not differ, (ii) that outbreak-associated EV strains like EV-D68 are circulating in South Africa and therefore ongoing surveillance is important, (iii) that E30 was the most prevalent EV strains and that E30 strains from respiratory illness cases in South Africa belongs to a newly proposed genotype k and lastly (iv) we conclude that rhinovirus species although found in asymptomatic illness cases may still contribute to a substantial number symptomatic illness cases across all age groups regardless of HIV status.

REFERENCES

1. Viral Meningitis [Internet]. Centres for Disease Control and Prevention. 2018. Available from: https://www.cdc.gov/meningitis/viral.html

2. Bailly JL, Mirand A, Henquell C, Archimbaud C, Chambon M, Charbonné F, et al. Phylogeography of circulating populations of human echovirus 30 over 50 years: Nucleotide polymorphism and signature of purifying selection in the VP1 capsid protein gene. Infect Genet Evol. 2009;9(4):699–708.

3. Pons-Salort M, Parker EPK, Grassly NC. The epidemiology of non-polio enteroviruses. Curr Opin Infect Dis [Internet]. 2015;28(5):479–87.

4. Dasaraju P V, Liu C. Chapter 93 Infections of the Respiratory System. In: Medical Microbiology. 1996.

5. GBD 2015 LRI Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect Dis. 2017 Nov;17(11):1133–61.

6. Gray DM, Zar HJ. Community-acquired pneumonia in HIV-infected children: a global perspective. Curr Opin Pulm Med. 2010 May;16(3):208–16.

7. Cohen C, Moyes J, Tempia S, Groom M, Walaza S, Pretorius M, et al. Severe influenza-associated respiratory infection in high HIV prevalence setting, South Africa, 2009-2011. Emerg Infect Dis. 2013;19(11):1766–74.

8. Tiveljung-Lindell A, Rotzén-Östlund M, Gupta S, Ullstrand R, Grillner L, Zweygberg-Wirgart B, et al. Development and implementation of a molecular diagnostic platform for daily rapid detection of 15 respiratory viruses. J Med Virol. 2009;81(1):167–75.

9. Brittain-Long R, Nord S, Olofsson S, Westin J, Anderson LM, Lindh M. Multiplex real-time PCR for detection of respiratory tract infections. J Clin Virol. 2008;41(1):53–6.

10. Pretorius MA, van Niekerk S, Tempia S, Moyes J, Cohen C, Madhi SA, et al. Replacement and positive evolution of subtype A and B respiratory syncytial virus G-protein genotypes from 1997-2012 in South Africa. J Infect Dis. 2013 Dec;208 Suppl:S227-37.

11. Cohen AL, Hellferscee O, Pretorius M, Treurnicht F, Walaza S, Madhi S, et al. Epidemiology of influenza virus types and subtypes in South Africa, 2009–20121. Emerg Infect Dis [Internet]. 2014;20(7):1149–56.

12. Haynes LM. Progress and challenges in RSV prophylaxis and vaccine development. J Infect Dis. 2013 Dec;208 Suppl:S177-83.

13. WHO. Pneumonia Fact sheet N°331 [Internet]. web page WHO. 2016. Available from: http://www.who.int/mediacentre/factsheets/fs331/en/

14. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ, et al. ICTV Virus Taxonomy Profile: Picornaviridae. J Gen Virol. 2017 Oct;98(10):2421–2.

15. Fields BN, Knipe DM, Howley PM. Fields Virology, 5th Edition [Internet]. Vol. 2, Fields Virology. 2007. 3177 p.

16. Alexander He, Koch G, Mountain Im, Sprunt K, Van Damme O. Infectivity of ribonucleic acid of poliovirus on HeLa cell mono-layers. Virology. 1958 Feb;5(1):172–3.

17. Hober D, Sauter P. Pathogenesis of type 1 diabetes mellitus: Interplay between enterovirus and host. Nature Reviews Endocrinology. 2010 May;6(5):279-89.

18. Oberste MS, Maher K, Kilpatrick DR, Flemister MR, Brown BA, Pallansch MA. Typing of human enteroviruses by partial sequencing of VP1. J Clin Microbiol. 1999;37(5):1288–93.

19. Palmenberg AC, Spiro D, Kuzmickas R, Wang S, Djikeng A, Rathe JA, et al. Sequencing and analyses of all known human rhinovirus genomes reveal structure and evolution. Science 2009 Apr 3;324(5923):55-9.

20. McIntyre CL, Knowles NJ, Simmonds P. Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. J Gen Virol. 2013 Aug;94(Pt 8):1791-806

21. Pretorius MA, Tempia S, Walaza S, Cohen AL, Moyes J, Variava E, et al. The role of influenza, RSV and other common respiratory viruses in severe acute respiratory infections and influenza-like illness in a population with a high HIV sero-prevalence, South Africa 2012-2015. J Clin Virol. 2016;75:21–6.

22. Carrillo-Tripp M, Shepherd CM, Borelli IA, Venkataraman S, Lander G, Natarajan P, et al. VIPERdb2: an enhanced and web API enabled relational database for structural virology. Nucleic Acids Res. 2009 Jan;37(Database issue):D436-42.

23. Jartti T, Lehtinen P, Vuorinen T, Koskenvuo M, Ruuskanen O. Persistence of Rhinovirus and Enterovirus RNA after Acute Respiratory Illness in Children. J Med Virol. 2004 Apr;72(4):695-9

24. Smura T, Ylipaasto P, Klemola P, Kaijalainen S, Kyllönen L, Sordi V, et al. Cellular tropism of human enterovirus D species serotypes EV-94, EV-70, and EV-68 in vitro: Implications for pathogenesis. J Med Virol. 2010;82(11):1940–9.

25. Wen H, Si L, Yuan X, Hao S, Gao F, Chu F, et al. Complete genome sequencing and analysis of six enterovirus 71 strains with different clinical phenotypes. Virol J [Internet]. 2013;10:115.

26. Pretorius MA, Madhi SA, Cohen C, Naidoo D, Groome M, Moyes J, et al. Respiratory viral coinfections identified by a 10-Plex real-time reverse-transcription polymerase chain reaction assay in patients hospitalized with severe acute respiratory illness-South Africa, 2009-2010. J Infect Dis. 2012

Dec 15;206 Suppl 1:S159-65..

27. Van Leer-Buter CC, Poelman R, Borger R, Niesters HGM. Newly identified Enterovirus C genotypes, identified in the Netherlands through routine sequencing of all Enteroviruses detected in clinical materials from 2008 to 2015.. J Clin Microbiol. 2016 Sep;54(9):2306-14

28. Tokarz R, Hirschberg DL, Sameroff S, Haq S, Luna G, Bennett AJ, et al. Genomic analysis of two novel human enterovirus C genotypes found in respiratory samples from Peru. J Gen Virol. 2013 Jan;94(Pt 1):120-7.

29. Renois F, Bouin A, Andreoletti L. Enterovirus 68 in pediatric patients hospitalized for acute airway diseases. J Clin Microbiol. 2013;51(2):640–3.

30. Xiang Z, Xie Z, Wang Z, Ren L, Xiao Y, Li L, et al. Human enterovirus genotype C104, China. Emerg Infect Dis. 2013 Apr;19(4):689-91

31. Xiang Z, Tsatsral S, Liu C, Li L, Ren L, Xiao Y, et al. Respiratory infection with enterovirus genotype C117, China and Mongolia. Emerg Infect Dis. 2014 Jun;20(6):1075-7

32. Puenpa J, Mauleekoonphairoj J, Linsuwanon P, Suwannakarn K, Chieochansin T, Korkong S, et al. Prevalence and characterization of enterovirus infections among pediatric patients with hand foot mouth disease, herpangina and influenza like illness in Thailand, 2012. PLoS One. 2014 Jun 2;9(6):e98888

33. Owens B. Rare enterovirus continues to circulate in North America. Vol. 384, Lancet (London, England). England; 2014. p. 1250.

34. Tokarz R, Firth C, Madhi SA, Howie SRC, Wu W, Sall AA, et al. Worldwide emergence of multiple clades of enterovirus 68. J Gen Virol. 2012;93(PART 9):1952–8.

35. Böttcher S, Prifert C, Weißbrich B, Adams O, Aldabbagh S, Eis-Hübinger AM, et al. Detection of enterovirus D68 in patients hospitalised in three tertiary university hospitals in Germany, 2013 to 2014. Eurosurveillance.

2016;21(19).

36. Gong YN, Yang SL, Shih SR, Huang YC, Chang PY, Huang CG, et al. Molecular evolution and the global reemergence of enterovirus D68 by genome-wide analysis. Med (United States). 2016 Aug;95(31):e4416

37. Imamura T, Fuji N, Suzuki A, Tamaki R, Saito M, Aniceto R, et al. Enterovirus 68 among children with severe acute respiratory infection, the Philippines. Emerg Infect Dis. 2011 Aug;17(8):1430-5

38. Henquell C, Mirand A, Richter J, Schuffenecker I, Böttiger B, Diedrich S, et al. Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission. J Virol [Internet]. 2013;87(22):12249–59.

39. Khetsuriani N, Lamonte-Fowlkes A, Oberst S, Pallansch MA. Enterovirus surveillance--United States, 1970-2005. MMWR Surveill Summ Morb Mortal Wkly report Surveill Summ / CDC. 2006;55(8):1–20.

40. Kyriakopoulou Z, Pliaka V, Amoutzias GD, Markoulatos P. Recombination among human non-polio enteroviruses: implications for epidemiology and evolution. Virus Genes. 2015 Apr;50(2):177–88.

41. Mertz D, Alawfi A, Pernica JM, Rutherford C, Luinstra K, Smieja M. Clinical severity of pediatric respiratory illness with enterovirus D68 compared with rhinovirus or other enterovirus genotypes. CMAJ. 2015 Nov;187(17):1279–84.

42. Harvala H, McIntyre CL, McLeish NJ, Kondracka J, Palmer J, Molyneaux P, et al. High detection frequency and viral loads of human rhinovirus species A to C in fecal samples; diagnostic and clinical implications. J Med Virol. 2012 Mar;84(3):536–42.

43. Royston L, Tapparel C. Rhinoviruses and Respiratory Enteroviruses: Not as Simple as ABC. Viruses. 2016 Jan;8(1).

44. Tapparel C, Junier T, Gerlach D, Van Belle S, Turin L, Cordey S, et al. New respiratory enterovirus and recombinant rhinoviruses among circulating picornaviruses. Emerg Infect Dis. 2009 May;15(5):719-26

45. Miller EK, Lu X, Erdman DD, Poehling KA, Zhu Y, Griffin MR, et al. Rhinovirus-associated hospitalizations in young children. J Infect Dis. 2007 Mar 15;195(6):773-81

46. Arden KE, McErlean P, Nissen MD, Sloots TP, Mackay IM. Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. J Med Virol. 2006 Sep;78(9):1232-40.

47. Fry AM, Lu X, Olsen SJ, Chittaganpitch M, Sawatwong P, Chantra S, et al. Human rhinovirus infections in rural Thailand: epidemiological evidence for rhinovirus as both pathogen and bystander. PLoS One. 2011 Mar;6(3):e17780.

48. Hammitt LL, Kazungu S, Morpeth SC, Gibson DG, Mvera B, Brent AJ, et al. A preliminary study of pneumonia etiology among hospitalized children in Kenya. Clin Infect Dis [Internet]. 2012 Apr;54 Suppl 2:S190-9.

49. Benet T, Sylla M, Messaoudi M, Sanchez Picot V, Telles J-N, Diakite A-A, et al. Etiology and factors associated with pneumonia in children under 5 years of age in mali: A Prospective case-control study. PLoS One. 2015;10(12):e0145447.

50. Breiman RF, Cosmas L, Njenga M, Williamson J, Mott JA, Katz MA, et al. Severe acute respiratory infection in children in a densely populated urban slum in Kenya, 2007-2011. BMC Infect Dis. 2015 Feb;15:95.

51. Self WH, Williams DJ, Zhu Y, Ampofo K, Pavia AT, Chappell JD, et al. Respiratory viral detection in children and adults: comparing asymptomatic controls and patients with community-acquired pneumonia. J Infect Dis. 2016 Feb;213(4):584–91.

52. Singleton RJ, Bulkow LR, Miernyk K, DeByle C, Pruitt L, Hummel KB, et al. Viral respiratory infections in hospitalized and community control children in Alaska. J Med Virol. 2010;82(7):1282–90.

53. Wolter N, Cohen C, Tempia S, Madhi SA, Venter M, Moyes J, et al. HIV and influenza virus infections are associated with increased blood pneumococcal load: a prospective, hospital-based observational study in South Africa, 2009-2011. J Infect Dis. 2014 Jan;209(1):56–65.

54. Tempia S, Walaza S, Viboud C, Cohen AL, Madhi SA, Venter M, et al. Mortality associated with seasonal and pandemic influenza and respiratory syncytial virus among children <5 years of age in a high HIV prevalence setting--South Africa, 1998-2009. Clin Infect Dis. 2014 May;58(9):1241–9.

55. Tempia S, Walaza S, Viboud C, Cohen AL, Madhi SA, Venter M, et al. Deaths associated with respiratory syncytial and influenza viruses among persons >/=5 years of age in HIV-prevalent area, South Africa, 1998-2009(1). Emerg Infect Dis. 2015 Apr;21(4):600–8.

56. Rames E, Macdonald J. Evaluation of MinION nanopore sequencing for rapid enterovirus genotyping. Virus Res. 2018 May;252:8–12.

57. Goya S, Valinotto LE, Tittarelli E, Rojo GL, Nabaes Jodar MS, Greninger AL, et al. An optimized methodology for whole genome sequencing of RNA respiratory viruses from nasopharyngeal aspirates. PLoS One. 2018 Jun 25;13(6):e0199714

58. Greninger AL. A decade of RNA virus metagenomics is (not) enough. Virus Res. 2018 Jan 15;244:218-229

59. Oberste MS, Maher K, Pallansch MA. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. J Virol. 2004 Jan;78(2):855-67.

60. Santti J, Hyypiä T, Kinnunen L, Salminen M. Evidence of recombination among enteroviruses. J Virol. 1999 Oct;73(10):8741-9.

61. Tan L Van, Tuyen NTK, Thanh TT, Ngan TT, Van HMT, Sabanathan S, et al. A generic assay for whole-genome amplification and deep sequencing of enterovirus A71. J Virol Methods. 2015 Apr;215–216:30–6.

62. Hirst GK. Genetic recombination with Newcastle disease virus, polioviruses, and influenza. Cold Spring Harb Symp Quant Biol. 1962;27:303-9.

63. Messacar K, Schreiner TL, Maloney JA, Wallace A, Ludke J, Oberste MS, et al. A cluster of acute flaccid paralysis and cranial nerve dysfunction temporally associated with an outbreak of enterovirus D68 in children in Colorado, USA. Lancet (London, England). 2015 Apr;385(9978):1662–71.

64. Allan Nix W, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. J Clin Microbiol. 2006;44(8):2698–704.

65. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

66. Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: More models, new heuristics and parallel computing. Nature Methods. 2012 Jul 30;9(8):772

67. Guindon S, Gascuel O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst Biol. 2003 Oct;52(5):696-704.

68. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24(8):1596–9.

69. Hall BG. Building phylogenetic trees from molecular data with MEGA. Mol Biol Evol. 2013;30(5):1229–35.

70. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, et al. Viral genome sequencing by random priming methods. BMC Genomics. 2008 Jan;9:5-2164-9–5.

71. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012 Jul 24;13:341.

72. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870-4.

73. Madden TL, Tatusov RL, Zhang J. Applications of network BLAST server. Methods in enzymology. 1996;266:131-41.

74. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015 May 26;1(1):vev003.

75. Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol. 2006 Oct;23(10):1891-901.

76. Delport W, Poon AF, Frost SD, Pond SLK. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. 2010 Oct;26(19):2455–7.

77. Zhou H, Alexander DH, Sehl ME, Sinsheimer JS, Sobel EM, Lange K. Penalized regression for genome-wide association screening of sequence data. Pac Symp Biocomput [Internet]. 2011;106–17.

78. Meijer A, Van Der Sanden S, Snijders BEP, Jaramillo-Gutierrez G, Bont L, Van Der Ent CK, et al. Emergence and epidemic occurrence of enterovirus 68 respiratory infections in The Netherlands in 2010. Virology [Internet]. 2012;423(1):49–57.

79. Renois F, Bouin A, Andreoletti L. Enterovirus 68 in pediatric patients hospitalized for acute airway diseases. J Clin Microbiol. 2013;51(2):640–3.

80. Meijer A, Benschop KS, Donker GA, van der Avoort HG. Continued seasonal circulation of enterovirus D68 in the Netherlands, 2011-2014. Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull. 2014;19(42).

81. Reiche J, Böttcher S, Diedrich S, Buchholz U, Buda S, Haas W, et al. Low-level circulation of enterovirus D68 associated acute respiratory infections, Germany, 2014. Emerg Infect Dis. 2015;21(5):837–41.

82. Tokarz R, Firth C, Madhi SA, Howie SRC, Wu W, Sall AA, et al. Worldwide emergence of multiple clades of enterovirus 68. J Gen Virol. 2012;93(PART 9):1952–8.

83. Du J, Zheng B, Zheng W, Li P, Kang J, Hou J, et al. Analysis of enterovirus 68 strains from the 2014 North American outbreak reveals a new clade, indicating viral evolution. PLoS One. 2015 Dec 2;10(12):e0144208.

84. Todd AK, Hall RJ, Wang J, Peacey M, McTavish S, Rand CJ, et al. Detection and whole genome sequence analysis of an enterovirus 68 cluster. Virol J. 2013 Apr 2;10:103.

85. Schuster JE, Miller JO, Selvarangan R, Weddle G, Thompson MT, Hassan F, et al. Severe enterovirus 68 respiratory illness in children requiring intensive care management. J Clin Virol. 2015 Sep;70:77-82.

86. Tapparel C, Siegrist F, Petty TJ, Kaiser L. Picornavirus and enterovirus diversity with associated human diseases. Infect Genet Evol. 2013 Mar;14:282-93.

87. Hellferscee O, Treurnicht FK, Tempia S, Variava E, Dawood H, Kahn K, et al. Enterovirus D68 and other enterovirus serotypes identified in South African patients with severe acute respiratory illness, 2009-2011. Influenza

Other Respi Viruses. 2017 May;11(3):211–9.

88. Yarmolskaya MS, Shumilina EY, Ivanova OE, Drexler JF, Lukashev AN. Molecular epidemiology of echoviruses 11 and 30 in Russia: Different properties of genotypes within an enterovirus serotype. Infect Genet Evol. 2015;30:244–8.

89. Liu N, Jia L, Yin J, Wu Z, Wang Z, Li P, Hao R et al. An outbreak of aseptic meningitis caused by a distinct lineage of coxsackievirus B5 in China. Int j Infect Dis [Internet]. 2014;23:101–4.

90. Papa A, Dumaidi K, Franzidou F, Antoniadis A. Genetic variation of coxsackie virus B5 strains associated with aseptic meningitis in Greece. Clin Microbiol Infect. 2006 Jul;12(7):688–91.

91. Midgley CM, Jackson MA, Selvarangan R, Turabelidze G, Obringer E, Johnson D, et al. Severe respiratory illness associated with enterovirus D68 - Missouri and Illinois, 2014. MMWR Morb Mortal Wkly Rep. 2014 Sep 12;63(36):798-9.

92. Patel MC, Wang W, Pletneva LM, Rajagopala S V, Tan Y, Hartert T V, et al. Enterovirus D-68 infection, prophylaxis, and vaccination in a novel permissive animal model, the Cotton Rat (Sigmodon hispidus). PLoS One. 2016;11(11):e0166336.

93. Janes VA, Minnaar R, Koen G, van Eijk H, Dijkman-de Haan K, Pajkrt D, et al. Presence of human non-polio enterovirus and parechovirus genotypes in an Amsterdam hospital in 2007 to 2011 compared to national and international published surveillance data: A comprehensive review. Vol. 19, Eurosurveillance. 2014. p. 1–9.

94. Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, et al. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. J Med Virol. 2017 Oct;89(10):1759–67.

95. Dyda A, Stelzer-Braid S, Adam D, Chughtai AA, Macintyre CR. The association between acute flaccid myelitis (AFM) and enterovirus D68 (EV-D68) – what is the evidence for causation? Eurosurveillance. 2018 Jan;23(3).

96. National Center for Biotechnology Information [Internet]. GenBank. 2018 [cited 2018 Jun 11]. Available from: https://www.ncbi.nlm.nih.gov/

97. Kyriakopoulou Z, Amoutzias GD, Dimitriou TG, Tsakogiannis D, Mossialos D, Markoulatos P. Intra- and inter-serotypic recombinations in the 5′ UTR-VP4 region of Echovirus 30 strains. Arch Virol. 2018 Feb;163(2):365-375.

98. Dimitrov KM, Sharma P, Volkening JD, Goraichuk I V., Wajid A, Rehmani SF, et al. A robust and cost-effective approach to sequence and analyze complete genomes of small RNA viruses. Virol J. 2017 Apr 7;14(1):72.

99. Chrzastek K, Lee D hun, Smith D, Sharma P, Suarez DL, Pantin-Jackwood M, et al. Use of Sequence-Independent, Single-Primer-Amplification (SISPA) for rapid detection, identification, and characterization of avian RNA viruses. Virology. 2017 Sep;509:159-166.

100. Hellferscee O, Tempia S, Walaza S, Variava E, Dawood H, Wolter N, et al. Enterovirus genotypes among patients with severe acute respiratory illness, influenza-like illness, and asymptomatic individuals in South Africa, 2012-2014. J Med Virol. 2017;89(10):1759–67.

101. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J Virol Methods. 2014 Jan;195:194-204.

102. Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE, Ilonen J. Recombination in circulating enteroviruses. J Virol. 2003 Oct;77(19):10423-31.

103. McWilliam Leitch EC, Cabrerizo M, Cardosa J, Harvala H, Ivanova OE, Kroes ACM, et al. Evolutionary dynamics and temporal/geographical correlates of recombination in the human enterovirus echovirus types 9, 11, and 30. J Virol [Internet]. 2010/07/07. 2010 Sep;84(18):9292–300.

104. Pallansch MA, Roos R. Enteroviruses: Polioviruses, coxsackieviruses, echoviruses and new enteroviruses. In: Knipe DM, Howley MP, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE. editiors. Fields Virology. 5th ed. Philadelphia: Lippincott Williams & Wilkens; 2007. p. 839-93

105. Pretorius MA, Tempia S, Treurnicht FK, Walaza S, Cohen AL, Moyes J, et al. Genetic diversity and molecular epidemiology of human rhinoviruses in South Africa. Influenza Other Respi Viruses. 2014 Sep;8(5):567–73.

106. Onyango CO, Welch SR, Munywoki PK, Agoti CN, Bett A, Ngama M, et al. Molecular epidemiology of human rhinovirus infections in Kilifi, coastal Kenya. J Med Virol. 2012 May;84(5):823–31.

107. Tempia S, Walaza S, Moyes J, Cohen AL, von Mollendorf C, McMorrow ML, et al. Attributable fraction of influenza virus detection to mild and severe respiratory illnesses in HIV-infected and HIV-uninfected patients, South Africa, 2012-2016. Emerg Infect Dis. 2017 Jul;23(7):1124–32.