TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

BIOINFORMATICS OF HUMAN PICORNAVIRUS MOLECULAR EVOLUTION AND PATHOGENESIS

Eero Hietanen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1676 | MEDICA – ODONTOLOGICA | TURKU 2022

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

BIOINFORMATICS OF HUMAN PICORNAVIRUS MOLECULAR EVOLUTION AND PATHOGENESIS

Eero Hietanen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1676 | MEDICA – ODONTOLOGICA | TURKU 2022

University of Turku

Faculty of Medicine
Institute of Biomedicine
Virology
Turku Doctoral Programme of Molecular Medicine

Supervised by

Docent, Petri Susi
Institute of Biomedicine/Virology
University of Turku
Turku, Finland

Reviewed by

Docent, Tero Ahola
Department of Microbiology
Faculty of Agriculture and Forestry
University of Helsinki
Helsinki, Finland

Docent, Jussi Hepojoki
Department of Virology
Medicum
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Institute of Veterinary Pathology
Vetsuisse Faculty
University of Zürich
Zürich, Switzerland

Opponent

Professor, Vesa Hytönen
BioMediTech
Faculty of Medicine and Health Technology
University of Tampere
Tampere, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-9081-8 (PRINT)
ISBN 978-951-29-9082-5 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku, Finland 2022

To my family

UNIVERSITY OF TURKU
Faculty of Medicine
Institute of Biomedicine
Virology
EERO HIETANEN: Bioinformatics of Human Picornavirus Molecular Evolution and Pathogenesis
Doctoral Dissertation, 131 pp.
Turku Doctoral Programme of Molecular Medicine
November 2022

ABSTRACT

Picornaviruses are a family of small, non-enveloped, positive-sense single-stranded RNA viruses. Together they form group of viruses that in humans cause diseases such as poliomyelitis, encephalitis, hand-foot-and-mouth disease, myocarditis, febrile illnesses, gastrointestinal infections, aseptic meningitis, and severe CNS infections. Characteristic to RNA viruses, picornaviruses also exhibit a high mutation rate and frequent recombination, making them a challenging target pathogen for research.

The emphasis of this dissertation was to highlight the importance of approachable bioinformatics in basic picornavirus research. The dissertation was divided into three parts, each of which had a different research theme from the point of picornaviral research. Bioinformatics was applied to the areas of molecular evolution of picornaviruses, their use in virotherapy, picornavirus pathogenesis, and antibody development.

In studies I and IV, Rigvir®, marketed as an effective oncolytic virotherapy drug against melanoma, was compared against clinical echovirus 7 isolates through sequence and structural bioinformatic analyses, as well as through *in vitro* infection assays. The results showed that Rigvir® did not differ in its ability to infect and lyse cancer or healthy cells when compared to the clinical isolates, casting doubt on its effectiveness as a virotherapy agent. In study II, recombination and receptor binding sites of coxsackievirus A9 were analyzed. The results showed that recombination occurs frequently in the genomes of picornaviruses, and that the sequence of the main receptor binding site remains conserved, while other putative heparan sulfate binding sites can still be found from previously unreported picornavirus types. In studies III and V, monoclonal antibodies and scFv fragments were developed against human parechovirus (HPeV) A1 VP0 protein. The developed antibodies not only recognized a wide range of HPeV-A1 strains but also HPeV genotypes 2–6, with the scFv showing potential in research use. Bioinformatic analyses concluded that the common epitope between different HPeV types is a structurally conformational epitope, with a conserved flanking region.

KEYWORDS: Picornavirus, bioinformatics, evolution, pathogenesis

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Biolääketieteen laitos
Virologia
EERO HIETANEN: Ihmisten pikornavirusten molekyylievoluution ja patogeneesin bioinformatiikka
Väitöskirja, 131 s.
Molekyylilääketieteen tohtoriohjelma
Marraskuu 2022

TIIVISTELMÄ

Pikornavirukset ovat pieniä, vaipattomia, positiivijuosteisia yksinauhaisia RNA viruksia. Pikornavirukset muodostavat yhden suurimmista ihmisiä infektoivista virusheimoista, johon kuuluvat virukset aiheuttavat tauteja kuten poliota, aivotulehdusta, käsi-suu-jalkatautia, sydänlihastulehdusta, kuumetauteja, mahatien infektioita, aseptistä meningiittiä ja vakavia keskushermoston infektioita. RNA viruksille tyypillisesti, myös pikornaviruket mutatoituvat nopeasti ja osoittavat jatkuvaa rekombinaatiota, jotka tekevät niistä haasteellisia patogeenejä tutkimukseen.

Väitöskirjan yhtenä tarkoituksena oli painottaa helppomuotoisen bioinformatiikan käyttöä ja tärkeyttä pikornavirusten perustutkimuksessa. Väitöskirjan julkaisut jakautuivat kolmeen eri teemaan, jotka käsittelivät ihmisten pikornavirusten bioinformatiikkaa liittyen niiden molekyylievoluution, käyttöön viroterapiassa, patogeneesin, ja vasta-ainekehitykseen.

Ensimmäisessä ja neljännessä osatyössä onkolyyttistä Rigvir® lääkevirusta verrattiin kliinisiin echovirus 7 isolaatteihin, ja näytettiin, että Rigvir® ei merkittävästi eronnut infektiivisyydeltään tai genomiltaan kliinisistä isolaateista. Tulokset siten antavat syytä epäillä Rigvir®:in toimintaa turvallisena ja tehokkaana lääkeviruksena. Toisessa osatyössä tarkasteltiin coxsackievirus A9:n rekombinaatiota ja reseptorien käyttöä. Tulokset osoittivat, että pikornavirusten rekombinaatio on jatkuvaa, ja reseptorin sitoutumiskohdat ovat konservoituneita luonnossa. Samalla pikornaviruksista voidaan löytää uusia sitoutumiskohtia heparaanisulfaatille virustyypeistä, joissa niitä ei ole ennen kuvailtu. Kolmannessa ja viidennessä osatyössä kehitettiin monoklonaalisia vasta-aineita ja yksiketjuisia variaabeli fragmentti (scFv) vasta-aineosia ihmisten parechovirus (HPeV) A1 VP0 proteiinia vastaan. Kehitetyt vasta-aineet tunnistivat laajan kirjon HPeV-A1 viruksia, sekä HPeV genotyyppejä 2–6. scFv vasta-aineet, joilla on merkittävää käyttöä diagnostiikassa, puolestaan osoittivat samanlaista tunnistavuutta. Bioinformatiivinen analyysi osoitti, että HPeV tyyppejä yhdistävä sitoutumiskohta on konformationaalinen epitooppi, jota ympäröi täysin konservoituneet alueet.

AVAINSANAT: Pikornavirus, bioinformatiikka, evoluutio, patogeneesi

Table of Contents

Abbreviations

aa	Amino acid
AP	Alkaline phosphatase
BCA	Bicinchoninic acid assay
BSA	Bovine serum albumin
CAR	Coxsackievirus-adenovirus receptor
cDNA	Complementary deoxyribonucleic acid
CDR	Complementarity determining region
CPE	Cytopathic effect
cre	*cis*-acting replication element
CV	Coxsackievirus
DAF	Decay-accelerating factor
DDBJ	DNA Data Bank of Japan
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
ENA	European Nucleotide Archive
EV	Enterovirus
FBS	Fetal bovine serum
GISAID	Global Initiative on Sharing Avian Influenza Data
GST	Glutathione S-transferase
HPeV	Human parecehovirus
HS	Heparan sulfate
ICAM 1	Intercellular Adhesion Molecule 1
ICTV	International Committee on Taxonomy of Viruses
IFA	Immunofluorescence assay
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IMAC	Immobilized metal affinity chromatography
INSDC	International Nucleotide Sequence Database Collaboration
IRES	Internal ribosome entry site
ITAF	IRES transacting factors
kb	Kilobase

MAFFT	Multiple Alignment using Fast Fourier Transform
MCMC	Markov chain Monte Carlo
MEGA7	Molecular Evolutionary Genetics Analysis 7
ML	Maximum likelihood
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
nt	Nucleotide
ORF	Open reading frame
PBS	Phosphate-buffered saline
RdRp	RNA-dependent RNA polymerase
RGD	Arginine-glycine-aspartic acid
RNA	Ribonucleic acid
PCR	Polymerase chain reaction
RT-qPCR	Reverse transcription quantitative real-time PCR
RV	Rhinovirus
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electroforesis
TBS	Tris-buffered saline
TCID50	Median tissue culture infectious dose
MRCA	Most common recent ancestor
TRF	Time-resolved fluorescence
UTR	Untranslated region
VPg	Viral protein genome-linked

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I **Hietanen, E.**, Smura, T., Hakanen, M., Chansaenroj, J., Merilahti, P., Nevalainen, J., Pandey, S., Koskinen, S., Tripathi, L., Poovorawan, Y., Pursiheimo, J., & Susi, P. (2018). Genome Sequences of RIGVIR Oncolytic Virotherapy Virus and Five Other Echovirus 7 Isolates. *Genome Announcements*, *6*(17), 7–8. doi: 10.1128/genomeA.00317-18

II **Hietanen, E.**, & Susi, P. (2020). Recombination events and conserved nature of receptor binding motifs in coxsackievirus A9 isolates. *Viruses*, *12*(1). doi: 10.3390/v12010068

III Tripathi, L., **Hietanen, E.**, Merilahti, P., Teixido, L., Sanchez-Alberola, N., Tauriainen, S., & Susi, P. (2021). Monoclonal antibody against VP0 recognizes a broad range of human parechoviruses. *Journal of Virological Methods*, *293*, 114167. doi: 10.1016/j.jviromet.2021.114167

IV **Hietanen, E.**, Koivu, M. K. A. A., & Susi, P. (2022). Cytolytic Properties and Genome Analysis of Rigvir® Oncolytic Virotherapy Virus and Other Echovirus 7 Isolates. *Viruses*, *14*(3), 525. doi: 10.3390/v14030525

V **Hietanen, E.***, Tripathi, L.*, Brockmann, E.-C., Merilahti, P., Lamminmäki, U., & Susi, P. (2022). Isolation and characterization of phage display-derived scFv antibodies against human parechovirus 1 VP0 protein. *Scientific Reports*, *12*(1), 13453. doi: 10.1038/s41598-022-17678-y (*equal contribution)

The original publications have been reproduced with the permission of the copyright holders.

1 Introduction

Members of the family *Picornaviridae* are small positive-sense, single-stranded, RNA viruses. They are capable of infecting a wide range of vertebrate hosts, including humans. Poliovirus, which belongs to the genus *Enterovirus*, is perhaps the best-known member of picornaviruses. Picornaviruses infecting humans form a diverse group of pathogens found in the *Cardiovirus*, *Cosavirus*, *Parechovirus*, *Kobuvirus*, *Salivirus*, and *Hepatovirus* genera (Cifuente & Moratorio, 2019; Nielsen et al., 2013). Consequently, these viruses rely on different infection routes, with varying epidemiology and pathogenesis. For example, enteroviruses infect the gastrointestinal tract, while rhinoviruses cause infections of the respiratory tract. Furthermore, the range of diseases that these viruses cause is vast, including, e.g., poliomyelitis, aseptic meningitis, respiratory illnesses like the common cold, hepatitis, encephalitis, hand-foot-and-mouth disease, and myocarditis (Pons-Salort et al., 2015). Life-threatening infections occur especially in neonates, making them a particular risk group of human-infecting picornaviruses (Lee et al., 2012; Shee & Weber, 2017). Furthermore, around the world, picornaviruses cause frequent epidemics by that threaten both human and agricultural health (Wolthers et al., 2019). While the impact of picornavirus infections on human health is immense, the specific infection mechanisms of numerous human picornaviruses remain unknown, which hinders the development of new therapeutic measures.

The family *Picornaviridae* is expanding rapidly, as well as the number of human-infecting virus species within the family. The technological advances in the biological sciences have enabled the sequencing and typing of more and more viruses with greater accuracy. The same technological advancements that have allowed us to generate the enormous amounts of data we see today also gives us to the tools to harness computational power to solve questions in biology. Bioinformatics combines computational methods with biology to analyse results of complex biological experiments and to guide research by offering predictions of new potential targets through data analysis and modelling. To gain understanding of a complex phenomenon such as human picornavirus infection, embracing the tools and methods provided by bioinformatics is critical because the amount of available data is colossal. To make use of bioinformatics, it is crucial that the future

development is considered from the viewpoints of both the biologists and the bioinformaticians. A dialogue and understanding of the methods between these disciplines is crucial in order to make bioinformatics more available to serve the underlying biological research better.

2 Review of the Literature

2.1 Picornavirus overview

2.1.1 Classification

Picornaviruses are non-enveloped single-stranded positive-sense RNA viruses that infect a wide range of different vertebrate hosts, including humans, and belong to the family *Picornaviridae*. Currently, the family *Picornaviridae* consists of 158 different species that are further grouped into 68 genera (Zell et al., 2017). The rapid growth of the family *Picornaviridae* is highlighted by the fact that since 2017, it has roughly doubled. The number of known genera has grown from 35 to 68, and the number of known species from 80 to 158 (Zell, 2018). The original classification of these viruses was reliant on the observed pathogenicity in animals (Stanway et al., 2014). However, due to the advancements of modern sequencing and computational methods in taxonomy, the current approach is to classify picornaviruses based on their genomic relationships (Stanway et al., 2014; Zell et al., 2017). While this change will arguably lead to more precise virus classification, it also means that the picornavirus taxonomy is currently being often changed as viruses are re-classified to their new taxa, which can lead to confusion. While the genetic diversity within the virus family is considerable, picornaviruses share distinct common features, and are further divided into sub-families, genera, and species based on their taxonomic classification derived from the genomic data that follows the criteria set for each species as determined by the International Committee on Taxonomy of Viruses (ICTV) (Zell et al., 2017). The vast majority of human-infecting picornaviruses belong to the *Enterovirus* genus, while other human infection picornaviruses can be found in the *Cardiovirus*, *Cosavirus*, *Parechovirus*, *Kobuvirus*, *Salivirus*, and *Hepatovirus* genera (Tapparel et al., 2013). This thesis work focuses on human-infecting picornaviruses, and particularly on viruses belonging to the *Enterovirus* and *Parechovirus* genera.

2.1.2 Picornavirus structure

The picornavirus virion consists of a non-enveloped capsid that protects the single-stranded viral RNA (Figure 1) (Zell et al., 2017). The virion size varies between 30 to 32 nm in diameter, and typically has relatively even structure surface. The structural motif uniting different picornaviruses is seen in the primary outward facing capsid proteins VP1, VP2, and VP3, which all have a common structure architecture consisting of eight-stranded β-barrels. Together the barrels form a tertiary structure that leads to icosahedral pseudo T-3 symmetry of the mature virion. A notable feature seen in enteroviruses is the canyon that forms around the fivefold axis, and is often associated with receptor binding (Baggen et al., 2018; Zell et al., 2017).

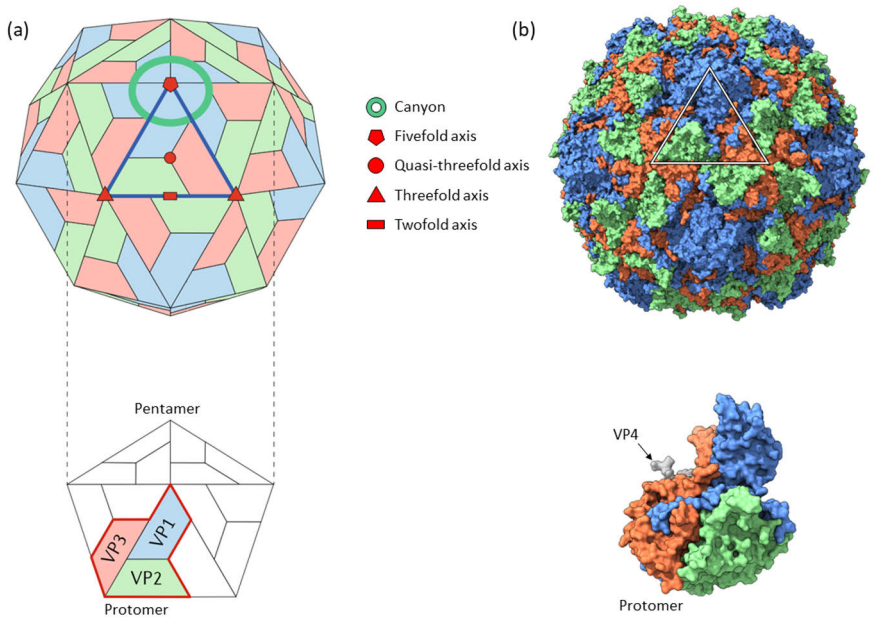

Figure 1. **(a)** Representative picornavirus capsid structure showing the protein organization and the canyon feature of the virion. **(b)** Poliovirus structure (PDB ID: 2PLV) shown from the same orientation highlighting the canyon feature depth on the surface. VP4 is seen below the protomer, positoning on the inner surface of the virion. Modified from Baggen at al. 2018.

2.1.3 Picornavirus genome

As compared to most RNA viruses, the genome of picornaviruses is quite small, typically ranging between 6.7 and 10.1 kb (Figure 2). The typical genome is divided into three regions: P1, P2, and P3, of which P1 encodes the viral capsid proteins,

while the P2 and P3 regions together encode for the non-structural proteins. The genome is translated into a single continuous polyprotein and further processed at specific cleavage sites between the regions, yielding functional structural and non-structural proteins (Zell et al., 2017). Additional cleavages generate mature functional proteins, or precursor proteins that serve different functions during the replication cycle before reaching mature forms following further cleavages (Lin et al., 2009; Zell et al., 2017).

The P1 genome region contains all of the structural proteins, 1A, 1B, 1C, and 1D, also known as the viral capsid proteins VP4, VP2, VP3, VP1 that form the capsid structure. The cleavage of precursor protein 1AB to yield VP4 and VP2 is hypothesized to take place during virion maturation. However, the cleaving of 1AB does not occur in all picornaviruses, and thus these viruses have only three capsid proteins: VP0, VP3, and VP1. In addition, many picornaviruses, such as the human-infecting viruses belonging to *Kobuvirus*, *Cardiovirus*, or *Salivirus* genera, possess an additional small leader protein, L, preceding the P1 region, which is a proteinase thought to be involved in later polyprotein processing (Lin et al., 2009). The 5′-UTR of picornaviruses varies between 0.5 to 1.5 kb in size, and includes many secondary RNA structures and domains that are important in the viral life cycle. Known structures of these include, for example, the poliovirus cloverleaf site, and an internal ribosomal entry site (IRES). The 3′-UTR contains a poly(A) tail, varies between 80 to 800 bases in length, and is also thought to contain secondary structures, such as pseudoknots that act as regulatory elements or have other functions in the viral life cycle (Francisco-Velilla et al., 2022; Martínez-Salas et al., 2015).

The P2 genome region contains 2A, 2B, and 2C proteins, which serve different functions during the replication cycle. 2A is a proteinase that takes part in the cleavage of the translated polyprotein, as well as has a role in disrupting translation functions of the host (Lin et al., 2009). 2B is assumed to take part in altering the host cell membrane permeability, in the formation of replication organelles, and in promoting virion release through viroporin activity (Li et al., 2019; Lin et al., 2009). 2C is a helicase that acts in junction with 2B to take part in the previously described functions of 2B (Porter, 1993).

The P3 genome region of the picornavirus genome encodes 3A, 3B, 3C, and 3D proteins. While the 3A membrane binding protein is not conserved at a sequence level across the picornavirus genera, the functionality of the protein appears to be conserved (Jackson & Belsham, 2021). The 3B protein is known as VPg (viral protein genome-linked), and is covalently attached to the 5′-UTR of the picornavirus genome, and acts as a primer for RNA synthesis (Lin et al., 2009; Zell et al., 2017). While the 3A protein is present in infected cells, it is more often found together with VPg as a part of the precursor protein 3AB. The precursor protein takes part in VPg uridylylation together with 3D, and it also mediates anchoring of the replication

complexes to the vesicles that later release the virus (Jackson & Belsham, 2021; Lin et al., 2009; Zell et al., 2017). The 3C protein is a proteinase that performs majority of the polyprotein processing and additionally cleaves host proteins in order to negatively affect host translation and to expedite the virus' own replication.The 3D protein is the viral RNA-dependent RNA polymerase (RdRp), which is the major complex responsible for the synthesis of new RNA during the replication cycle. The 3C and 3D proteins also form a precursor protein, 3CD, which exhibits suppressed proteolytic activity that is released following release and maturation of the individual proteins (Peersen, 2017; Winston & Boehr, 2021). The 3CD precursor protein serves important functions, such as regulating replication through modulating protease domains and interacting with replication membranes (Peersen, 2017; Winston & Boehr, 2021; Xiang et al., 1995).

The picornaviral genomes also harbour an important internal regulatory site known as the *cis*-acting replication element (*cre*), however, its location varies between the genera of the family. The *cre* is an RNA stem-loop structure that is involved in the initiation of replication together with the VPg, and the viral RdRp (Paul & Wimmer, 2015; Steil & Barton, 2009).

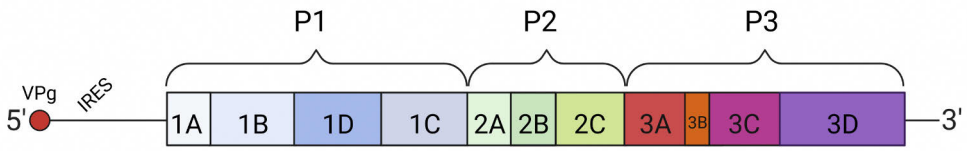

Figure 2. Layout of a typical picornavirus genome.

2.2 Enteroviruses

The *Enterovirus* genus divides into 15 different species, namely *Enterovirus A – L* and *Rhinovirus A – C*. Enteroviruses infecting humans are found in the species *Enterovirus A – D* and *Rhinovirus A – C*. The type species of the genus is *Enterovirus C*, which includes the well-known polioviruses. Enterovirus genome varies between 7100 and 7450 bases in length, encoding a polyprotein that is approximately between 2100 to 2200 amino acids in length. The genome layout of human-infecting enteroviruses follows the generic layout of the virus family (Figure 2) and contains a type I IRES. The location of the *cre* in enteroviruses changes between species, and can be found from the genomic regions corresponding to the 2C, 2A, 1D, or 1B proteins. The virions of entero- and rhinoviruses are larger compared to an average picornavirus, at roughly 30 nm in size. Factors contributing to the larger virion size include the longer 1B, 1C, and 1D capsid protein lengths, in turn causing longer loops connecting the capsid proteins. Consequently, this leads to a thicker capsid

wall, and more rugged surface topology compared to other picornaviruses due to the deep "canyon" feature surrounding VP1 in enteroviruses (Tuthill et al., 2010; Zell et al., 2017). The species demarcation criteria as defined by the ICTV state that the members of the *Enterovirus* genus are of the same species, if they are less than 30% divergent in the polyprotein amino acid sequence, less than 40% divergent in the P1 region amino acid sequence, less than 30% divergent in the combined amino acid sequence of the 2C and 3CD sequence, and share a common genome organization (Zell et al., 2017).

Enteroviruses mainly infect the gastrointestinal tract, have a fecal-oral transmission route, and are particularly frequently seen infecting young children (Muehlenbachs et al., 2015). Rhinoviruses infect the upper respiratory tract and are transmitted through respiratory droplets. Rhinoviruses are the major cause of the common cold, infecting both children and adults alike. While more severe rhinovirus infections do not occur frequently, the burden on healthcare is still significant due to the widespread of rhinovirus infections (Tapparel et al., 2013). On the other hand, enterovirus infections are typically asymptomatic, but they can present in severe and life-threatening forms. The high diversity of enterovirus species provides an additional challenge to the development of effective therapies against enterovirus infection. Enterovirus infections present in a wide variety of different phenotypes, some of which include poliomyelitis, encephalitis, herpangina, hand-foot-and-mouth disease, myocarditis, febrile illnesses, conjunctivitis, hepatitis, aseptic meningitis, and respiratory illnesses (De Crom et al., 2016; Tapparel et al., 2013). Severity and type of clinical symptoms of enterovirus infection are also age. Especially neonates and infants often suffer severe infections with symptoms including myocarditis, central nervous system (CNS) diseases, and sepsis-like illnesses (De Crom et al., 2016; Muehlenbachs et al., 2015). While the polioviruses are a well-known cause of neurological symptoms, other enterovirus types, such as enterovirus 71 (EV-71) and viruses in the coxsackievirus A (CVA) group also cause infections that present neurological symptoms (Muehlenbachs et al., 2015). Furthermore, enteroviruses contribute massively to the occurrence of viral meningitis, causing more than 90% of all cases. While enteroviruses present a considerable public health risk, their diversity provides a serious challenge for the development of working therapies.

2.3 Parechoviruses

The *Parechovirus* genus includes 6 different species, namely *Parechovirus A – F*. Known human-infecting parechoviruses are confined to the species *Parechovirus A*, which itself contains 19 different HPeV types. Regardless of the amount of parechovirus types included in the *A* species, research has focused primarily on HPeV types 1–3 (De Crom et al., 2016). The parechovirus genome length varies

between 7300 and 7600 bases, encoding a polyprotein of roughly between 2180 and 2250 amino acids in length. The genome includes a type II IRES, with the *cre* identified within the 1AB genome region of the HPeV genome (Zell et al., 2017). Contrary to enteroviruses, all parechoviruses contain a specific H-box/NC motif in the 2A protein, although the specific function of the highly conserved motif remains unknown (Hughes & Stanway, 2000; Zell, 2018). The parechovirus virion is slightly smaller than that of enteroviruses at approximately 30 nm in diameter, and also exhibits a smoother surface topology (Zell et al., 2017). The species demarcation criteria as defined by the ICTV is that the members of *Parechovirus* genus are of the same species, if they are less than 30% divergent in the polyprotein amino acid sequence, less than 30% divergent in the P1 region amino acid sequence, less than 20% divergent in the combined amino acid sequence of the non-structural proteins 2C and 3CD, and share a common genome organization (Zell et al., 2017).

Like enteroviruses, HPeVs infect the gastrointestinal and respiratory tract. Similarly, infections are particularly prevalent in infants and young children (De Crom et al., 2016; Shee & Weber, 2017; Sridhar et al., 2019). The wide prevalence of parechovirus infections is exemplified by studies showing that more than 90% of children under the age of 2 years old have been exposed to a parechovirus (De Crom et al., 2016; Joki-Korpela & Hyypiä, 1998; Tauriainen et al., 2007). As previously mentioned, research has focused primarily on HPeV types 1–3. HPeV-1 and HPeV-2 often cause asymptomatic infections. HPeV-1 is also the most commonly circulating type, causing gastrointestinal symptoms and respiratory illnesses (De Crom et al., 2016; Shee & Weber, 2017). HPeV-3 is of particular interest, as it causes considerably more severe infections and contributes significantly to the number of viral meningitis and sepsis-like illnesses (De Crom et al., 2016; Shee & Weber, 2017; Sridhar et al., 2019). Furthermore, HPeV-3 differs epidemiologically from the other types, as epidemics occur in clusters rather than sporadically throughout a year (Shee & Weber, 2017). Altogether, on top of their pre-existing burden on healthcare, HPeVs show dangerous prospects with the wide prevalence of the commonly circulating type HPeV-1, and the severe infection causing HPeV-3. The lack of a specific arginine-glycine-glutamic acid (RGD) motif found in the VP1 capsid protein of HPeV-1 but not in HPeV-3 has been suggested to contribute to severity of HPeV-3 infection (De Crom et al., 2016; Shee & Weber, 2017). The RGD motif is involved in integrin receptor binding in picornaviruses, and thus HPeV-3 is thought to instead use other receptors and to potentially have a different tissue tropism (Boonyakiat et al., 2001; De Crom et al., 2016; Mason et al., 1994; Shee & Weber, 2017; Sridhar et al., 2019).

2.4 Picornavirus infection cycle

Picornaviruses employ a larger number of different cellular receptors for cell entry, and the receptors are often species-dependent. However, not all receptors used by picornaviruses are known, and the cellular entry mechanisms of picornaviruses are thus not fully known (Tuthill et al., 2010). This section covers the general lifecycle of picornaviruses as they infect a host, gain cellular entry, replicate, and ultimately release from the host cell to start the cycle anew (Figure 3).

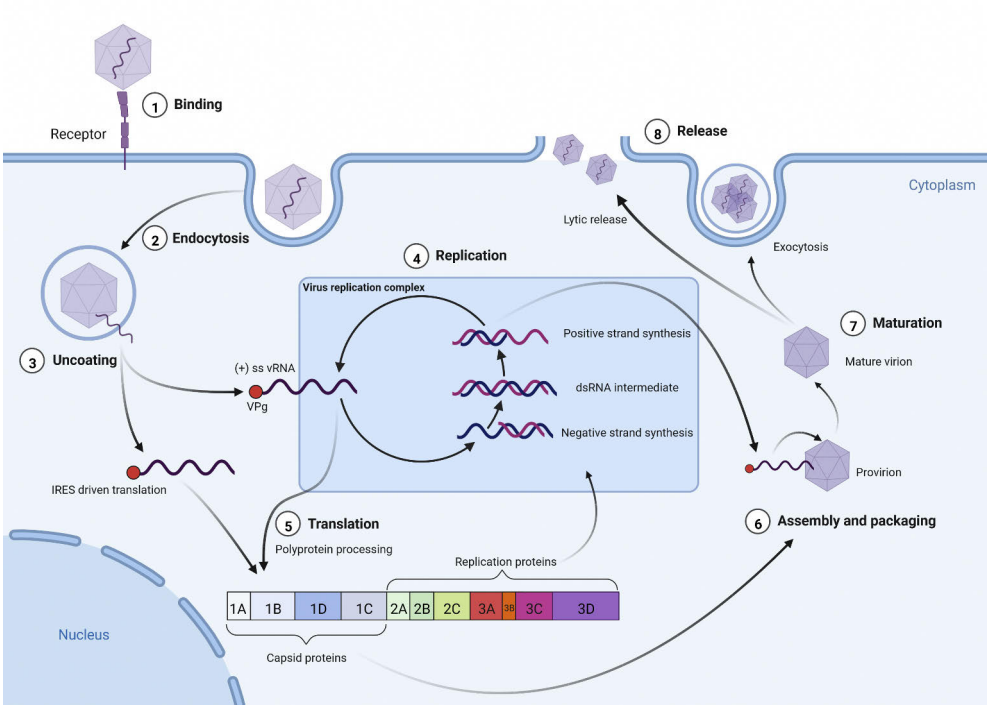

Figure 3. General picornavirus infection cycle.

At the cellular level, the picornavirus infection cycle starts when the virus recognizes a specific cell-surface receptor, or multiple receptors. Following receptor binding the virus is internalized through endocytosis (Baggen et al., 2018; Tuthill et al., 2010). While many picornaviruses lack conclusive evidence regarding the receptors they utilize during infection, some have been well established. For example, the intracellular adhesion molecule 1 (ICAM-1) receptor is utilized by the major group of human rhinoviruses, as well as coxsackieviruses A21 and A24 (Baggen et al., 2018; Shafren et al., 1997; Tuthill et al., 2010). The coxsackievirus-adenovirus receptor (CAR) acts as a target for many coxsackievirus B group viruses (Baggen et al., 2018; Martino et al., 2000; Tuthill et al., 2010). Integrins, such as

$\alpha_v\beta_3$ and $\alpha_v\beta_6$, have also been recognized receptors for many picornaviruses, including echovirus 1, echovirus 9, coxsackievirus A9, parechovirus A1 (Heikkilä et al., 2009; Joki-Korpela et al., 2001; M. Triantafilou et al., 2001). Additionally, the functionality of integrins acting as picornavirus receptors is often seen depend on the recognition of a specific RGD motif in the VP1 protein, although interestingly, it might not be necessary for successful infection in all integrin utilizing picornaviruses (Boonyakiat et al., 2001; Hughes et al., 1995; Ioannou & Stanway, 2021; Manning et al., 2008; Mason et al., 1994; Roivainen et al., 1991; Williams et al., 2004). Some integrin binding viruses have also been reported to be able to adapt to utilizing heparan sulfate (HS) as their receptor, which indicates a possible alternate RGD-independent entry route for these viruses (Baeshen, 2014; Jackson et al., 2000; McLeish et al., 2012; Tuthill et al., 2010). The decay-accelerating factor (DAF) on the other hand serves as a receptor for coxsackieviruses A21, B1, and B3, as well as many echoviruses (Baggen et al., 2018; Bergelson et al., 1994). Additionally, some viruses, such as those that utilize DAF as their primary receptor, may require co-receptors in addition to the primary receptor, in order to gain cellular entry (Tuthill et al., 2010). Interestingly, the primary receptor is not always necessary for successful infection, and the virus can infect the cell simply by exploiting the co-receptor. The disparity as to why some viruses can infect a cell simply by utilizing a molecule that is commonly thought to be a co-receptor is not fully clear (Tuthill et al., 2010). However, some studies suggest that the differences in conformational changes taking place in the capsid under native and laboratory conditions would be a contributing factor. Additionally, the cellular models used in studies have been proposed as a further explanation (Tuthill et al., 2010).

After endocytosis, uncoating of the virion releases the viral RNA (vRNA) into the cytoplasm. The detailed mechanisms of uncoating differ between picornaviruses, and the process is still poorly understood as a whole (Ren et al., 2013; Tuthill et al., 2010). Studies with picornaviruses whose structure have been resolved at high resolution suggest the uncoating to require a pocket factor, which is a surface structure within the VP1 capsid protein that is inhabited by a fatty acid element derived from the host cell during mature virion exocytosis (Smyth et al., 2003; Smyth & Martin, 2002). The release of the pocket factor from the capsid surface is thought to contribute to the destabilization of the virion, thus encouraging uncoating (Ren et al., 2013; Smyth et al., 2003; Smyth & Martin, 2002). Another key event in uncoating has been linked to the VP4 capsid protein and its loss from the virion, which is suggested to happen through interactions between VP4 and the host cell membranes (Smyth & Martin, 2002).

After the vRNA release into the host cytoplasm, the genome can go through cap-independent translation that is mediated by the IRES, and enabled by different translation initiation factors and RNA binding proteins called IRES transacting

factors (ITAFs) (Francisco-Velilla et al., 2022; Lin et al., 2009; Martínez-Salas et al., 2015). Through cap-independent translation and the involved polyprotein processing, the host cell is exposed to the non-structural viral proteins from the P2 and P3 genome regions, which are involved in shutting down the host cells translational machinery, formation of the replication complexes, and in the replication of new viral genomes (Baggen et al., 2018; Jackson & Belsham, 2021; Lin et al., 2009).

Further replication can take place in the membranous replication complexes formed from the host cells endoplasmic reticulum and membranes from the Golgi apparatus (Baggen et al., 2018; Belov, 2016). Replication within the complex is reliant on the RdRp. As single-stranded positive sense RNA viruses, the replication process within the complex proceeds with the synthesis of the negative-strand RNA that in turn acts as a template for new positive-strand synthesis (Baggen et al., 2018). Subsequently, the newly synthesized RNA can act as a template for following replication rounds, or they can be egressed from the replication complex to go through translation, or be packaged into the forming provirions (Baggen et al., 2018). As mentioned above, the translation is mediated by the IRES, and processing of the polyprotein into replication complex subunits further accelerates the replication process. Cleaved structural proteins form protomers that eventually self-assemble into provirion structures, and into which the viral genomes are packaged. The provirion assembly is not fully understood, but it involves many of the viral proteins and various host proteins (Baggen et al., 2018; Lin et al., 2009).

Last step of the infection cycle involves the maturation of the provirion, which occurs through VP0 cleavage to yield VP2 and VP4 as catalyzed by the genomic RNA within the provirion (Baggen et al., 2018). Most often the mature virions exit the cell through a lytic route, which involved the destruction of the host cell (Baggen et al., 2018; Zell et al., 2017). However, the mature virions can also be released non-lytically within vesicles that are formed through the triggering of autophagy. This process involves the formation of autophagosomes that transfer clusters of mature virions out of the host cell (Baggen et al., 2018; Chen et al., 2015; Lai et al., 2016).

2.5 Picornavirus evolution

Like other RNA-viruses, picornaviruses have a high mutation rate, which is primarily due to the low accuracy and lack of error correction mechanisms in the RdRp, as opposed to DNA polymerases (Agol, 2014; Ferrer-Orta et al., 2015; Puenpa et al., 2016). The above leads inevitably to faster evolution of the virus' genome. Considering the amount of genome copies produced during a viral infection, this fast-paced evolution makes the research of RNA viruses problematic especially when the goal is to develop effective drugs against viral infection.

2.5.1 Quasispecies concept

Quasispecies is a term meaning a group of mutants, deriving from a system that is under selection pressure, and acts on the species as a whole. While the quasispecies concept can be applied to any species, it has become especially useful when considering RNA viruses, as the effects of rapid evolution can be seen more clearly in them. In the case of RNA viruses, the term quasispecies can be used to refer to a single pool of mutant genomes that arises from viral infection (Domingo & Perales, 2019). The force generating the swarm of RNA genomes is the RdRp that directly contributes to the high number of newly generated mutant genomes, even if the template can be thought to originate from a single "master" genome. Furthermore, as the viral infection proceeds and newly generated genomes are used as new templates for continued replication, the effects of the inaccurately acting RdRp are compounded further (Peersen, 2017). Additionally, other effects such as recombination and the cellular environment add their own effects on the viral replication process. As a result, although it could be considered that a viral infection can start from a single particle with a specific genome, the result is that the released virions carrying out the infection are not representative of the originally infecting genome. As such, the virus can be thought to live as a "swarm" or a "cloud" of mutant genomes that have direct implications to the pathogenicity and tropism of the virus (Domingo & Perales, 2019; Sanjuán & Domingo-Calap, 2016).

2.5.2 Mutations

The types of mutations that can occur during replication can be categorized into deletions, insertions, and base substitutions. Deletions remove nucleotides from the sequence, insertions add them, and substitutions exchange nucleotides to others. Insertions and deletions can lead to a non-functional protein product if the translation frame is shifted by other than a complete codon, or the result can be a non-functional protein due to the insertion or a deletion of an amino acid. Substitutions are also divided into either transitions or transversions. Transitions change either a pyrimidine or a purine to another pyrimidine or a purine, while transitions either change a pyrimidine to a purine, or a purine to a pyrimidine (Guo et al., 2017; Mahdieh & Rabbani, 2013).

Mutations occurring in a genome region that encodes for a functional protein can result in either a silent, a missense, or a nonsense mutation. A silent mutation, also called a synonymous substitution, refers to a nucleotide substitution in the codon that does not change the translated amino acid. A missense mutation is a change in a codon that results in a different amino acid being translated. Such mutations can potentially have a significant effect on the functionality of the protein product, as amino acid substitutions can have an effect on, for example, protein folding and

target binding, by drastically changing the chemical properties of the incorporated amino acid. Finally, a nonsense mutation results in a nucleotide change that causes a premature stop codon to appear in the genome, often resulting in a non-functional and truncated protein. Nonsense and missense mutations are also referred to as nonsynonymous substitutions (Mahdieh & Rabbani, 2013).

Mutations can thus have a multitude of effects on the viability of a virus through different mechanisms. Some mutations within the viral genome do not cause any significant change in the functionality of the virus, while others either hinder or enhance its effectiveness to infect new cells. Many human diseases have also been linked to a single point mutation in a specific gene, and likewise in viruses, single mutations have been linked to increased cellular tropism and antigenicity, highlighting the potentially drastic effect that a small change in the genome can cause (Baloch et al., 2021; Etingov et al., 2008; Shastry & Shastry, 2002; Wang & Moult, 2001; Yuan et al., 2021). Additionally, while mutations can be beneficial for the virus, there is also a naturally occurring upper limit that the genome can handle. A genome strained too high by deleterious mutations will eventually lead to loss of viability, which, in fact, can be exploited by artificially increasing the frequency of mutations through introduction of a mutagen, in a process referred to as lethal mutagenesis (Duffy, 2018; Martin & Gandon, 2010).

2.5.3 Recombination

Recombination refers to the process where two different genome templates are contributing to the replication of a newly synthesized genome. Recombination can be divided into recombination through genomic reassortment, and the more classical recombination of template switching. Reassortment can occur when the genome is divided into segments, as is the case with e.g. influenza virus, belonging to the *Orthomyxoviridae* family. In reassortment, whole genome segments are be interchanged, which leads to newly created genome combinations (Lowen, 2018; Simon-Loriere & Holmes, 2011). Reassortment is confined to viruses with segmented genomes. On the other hand, classic recombination is much more prevalent in RNA viruses, and is widely considered to be possible in most, if not all, of them at varying rates (Pérez-Losada et al., 2015; Simon-Loriere & Holmes, 2011). In classical recombination, as different viruses infect the same host, the templates used by the RdRp during the replication process can be switched sporadically (Simon-Loriere & Holmes, 2011). Additionally, this type of recombination can also occur as inter- or intraspecies, although in the case of picornaviruses, interspecies recombination is not as often seen in nature (Bentley & Evans, 2018; Furione et al., 1993; Santti et al., 1999, 2000; Smura et al., 2013).

As with reassortment, classical recombination can have profound effects on the phenotype of a virus. Genome level changes arising from recombination events can lead to very rapid changes in things such as immune response, cellular tropism, and pathogenicity (Bentley & Evans, 2018; Simon-Loriere & Holmes, 2011). For example, recombination events affecting the IRES elements located in the 5′-UTR of picornaviruses contribute to changes in tissue tropism. Additionally, the possibility of recombination across different virus genera and families can lead to novel IRES element recombinations that are a mix of IRES types I–IV (Arhab et al., 2020; Martínez-Salas et al., 2015).

Lastly, the study of recombination events has become an important tool in virology. As an example, more accurate virus classification can be achieved by studying recombination, and recombination events between viruses can in epidemiological studies help tracking the movement of viruses across geographical areas (Bentley & Evans, 2018; Lukashev, 2010; Simon-Loriere & Holmes, 2011; Sugauchi et al., 2003).

2.6 Bioinformatics

Although nucleotide sequencing and the handling of sequence data are nowadays most directly associated to the term bioinformatics, the field in fact emerged in the early 1960s, when it was first applied to protein sequence analysis (Gauthier et al., 2019). Later, new methods created for DNA analysis and the developments in computer science have enabled bioinformatics to be applied to a much wider range of biological research problems (Gauthier et al., 2019). Although the roots of bioinformatics now date back approximately 60 years, it can be argued that the most significant advancements in the field came with the development of new sequencing technologies and the vast increase in the computation power available (Gauthier et al., 2019).

2.6.1 Application of bioinformatics

The underlying methods in bioinformatics pull heavily from computer science, from where the development of algorithms and computational methods for modern bioinformatics originates. Of course, in the end bioinformatics is an amalgamation of many scientific fields, including computer science, medicine, biology, chemistry, mathematics, and physics (Bayat, 2002; Gauthier et al., 2019). While the development of bioinformatic methods for their own sake is important in order to advance the field, it is equally important that bioinformatics is utilized even in basic biological research, where it is in fact often even necessary due to the complicated nature of modern research. While bioinformaticians are equipped to develop custom analysis pipelines

to address specific research questions, or handle vast amounts of biological data for analysis, these methods are often out of reach for regular biologists whose expertise lies elsewhere (Hufsky et al., 2018). Modern biological research, however, often includes demanding data analysis, and bioinformatics, at least in some form. As such, it is important that bioinformaticians and biologists both take an interest in the type bioinformatics can also be used by those who are not necessarily well equipped with all of the theory and practical skills behind bioinformatics. This means that the development side of bioinformatics should take the end user into account, and build tools in such a way that they can be approached by biologists (Hufsky et al., 2018; Marz et al., 2014). At the same time, biologists should take an interest in utilizing bioinformatics in their own work, as outsourcing the analyses to bioinformaticians is not always an option. A mutual understanding of the respective fields is also important, so that bioinformaticians and biologists can share a common language when discussing bioinformatic needs for biological research.

Virology is a field that can especially benefit from the heavy use of bioinformatics. Some headway has been made in an effort to connect bioinformaticians with virologists to tackle specific research issues. For example, the foundation of the European Virus Bioinformatics Center established in 2016 aims to do exactly this (Hufsky et al., 2018).

2.6.2 Biological databases

Biological databases could be considered the backbone of bioinformatics. New data is generated at an ever-increasing speed, and biological databases have a central role in storing and organizing data so that it can be publicly available for everyone to use. While some databases have a more general role, such as GenBank (Sayers et al., 2021), in storing all manner of sequence data, some databases take a more specified role and serve as repositories for only a certain organism or type of data. While data can take several forms, it can be argued that the two most fundamental forms of data that enable bioinformatics are sequence and structural data.

2.6.2.1 Sequence data

Sequence data refers to the sequenced genomic information. Most commonly, this simply refers to information of a nucleic acid sequence consisting of letters depicting the nucleobases adenine, thymine (uracil in the case of RNA), guanine, or cytosine as A, T (U), G, and C for DNA. Nucleotide sequence data can be translated to a protein sequence automatically, and thus protein sequencing itself is not explicitly necessary. Additionally, direct amino acid sequencing is considerably more difficult compared to modern DNA sequencing methods as it requires more complex or slow

methods such as mass spectrometry or Edman degradation (Smith, 2001; Yates et al., 1995). Protein sequence data uses a specific code assigned for each amino acid, as generally agreed by the International Union of Pure and Applied Chemistry (IUPAC) (Johnson, 2010). While there is a standard codon table, the code can change between different organisms, who have been assigned their own codon codes for amino acid translation (Jukes & Osawa, 1993). The accuracy of the obtained sequence data depends on a multitude of factors, such as the quality of the sample, care taken during sample preparation, sequencing technology used, possible sequence assembly methods and quality control. Still, especially considering virology, an assembled consensus sequence is only a snapshot of a virus and its genome. Considerations of phenomenon such as the quasispecies, and what effects it might have on our understanding of a viral genome should be considered.

2.6.2.2 Structural data

Structural data refers to the data obtained often via e.g. X-ray crystallography. X-ray crystallography allows determining the structure of the target molecule at atomic level. In biological research, crystallography is used to e.g. determine the structure of proteins. More recently, cryo-EM has emerged as a significant method for resolving more accurate native conformations of virus particles (Jiang & Tang, 2017). In virological research the structural proteins are of particular interest with non-enveloped viruses such as picornaviruses, because their structures often have direct implications to pathogenicity. While sequence data is essentially a string of letters with some specific metadata added, the form of structural data is more complex and requires more sophisticated analysis tools to access the 3-dimensional organization of the target. However, structural data allows better understanding of the molecule's native state and the associated biological functions.

2.6.2.3 Common databases and central hubs

Table 1 presents the most prominent biological databases or hubs. The largest central hub for DNA sequence data is the International Nucleotide Sequence Database Collaboration (INSDC) (Arita et al., 2021). The INSDC is a collaboration between the DNA Data Bank of Japan (DDBJ) (Mashima et al., 2017), the European Nucleotide Archive (ENA) (Cummins et al., 2022), and GenBank (Sayers et al., 2021). INSDC and its collaborators offer sequence data in the form of raw next-generation sequencing reads, assembled data, sample data, and data relating to single research projects in singular packages. The collaborators of INSDC also offer multiple more specialized databases. For example, GenBank, run by the National Center for Biotechnology Information, contains numerous databases that include

access to data sets with more included metadata, or higher-level information, such as taxonomic data, protein data, and data sets from comparative studies.

Some more specific databases include for example UniProt, which specializes in protein data and additional information surrounding the function of proteins. Other specialized databases include the GISAID database, which is focused on sharing specifically influenza data. Other prominent viruses also, such as HIV, have their own database. Some pathogens, such as picornaviruses also have their own commonly recognized website that upholds information about the pathogen and links to external databases for further data access (https://www.picornaviridae.com/).

Table 1. Prominent biological databases, web, and standalone tools. (Modified from Phadke et al., 2021).

CATEGORY	NAME	TYPES OF DATA	WEBLINK
DATABASE	INSDC	Sequence data hub	https://www.insdc.org
	GenBank	Gene and genome sequences	https://www.ncbi.nlm.nih.gov/genbank
	DNA Data Bank of Japan	Gene and genome sequences	https://www.ddbj.nig.ac.jp
	European Nucleotide Archive	Gene and genome sequences	https://www.ebi.ac.uk/ena/browser
	UniProt	Protein sequences	https://www.uniprot.org
	RefSeq	Curated genome and protein sequences	https://www.ncbi.nlm.nih.gov/refseq
	Worldwide Protein Data Bank	3D protein structures	https://www.wwpdb.org
	Immune Epitope Database	Experimental data on B cell and T cell epitopes	https://www.iedb.org
	Virus Particle Explorer	Structures of viruses	https://viperdb.org
	Viral Pathogen Resource	Comprehensive data on high priority human pathogenic and related viruses	https://www.viprbrc.org
	GISAID	Comprehensive collection of influenza related data	https://www.gisaid.org

CATEGORY	NAME	TYPES OF DATA	WEBLINK
	FluNet	Global influenza surveillance network	https://www.who.int/tools/flunet
	Hepatitis B Virus Database	Sequence data and analysis of drug resistance profiling	https://hbvdb.lyon.inserm.fr
	HIV Database	Sequence data and immunological epitopes	https://www.hiv.lanl.gov/content/index
	ViralZone	Extensive and curated viral database	https://viralzone.expasy.org
	Protocols.io	Protocols for biological assays	https://www.protocols.io
TOOLS (WEB)	NextStrain	Rapid tracking of ongoing infectious outbreaks through phylogenetics	https://nextstrain.org
	Galaxy	Many tools for data based biological research	https://usegalaxy.org
	EDGE	Tools for NGS sequence assembly and analysis	https://edgebioinformatics.org
	EVBC tools collection	Many links to external tools for virological research	https://evbc.uni-jena.de
TOOLS (STANDALONE)	MEGA	Sequence analysis and phylogenetics	https://www.megasoftware.net
	Geneious	Comprehensive bioinformatics platform (commercial)	https://www.geneious.com
	BEAST	Comprehensive Bayesian phylogenetics software package	https://www.beast2.org
	UCSF Chimera	Platform for the analysis of structural data	https://www.cgl.ucsf.edu/chimera
	PyMOL	Platform for the analysis of structural data	https://pymol.org/2

2.6.3 Software

2.6.3.1 Problem of abundance, target audience, and mode of use

Some of the popular resources for bioinformatic tools are presented in Table 1. While bioinformatic software development is essential, the way in which bioinformatic software are developed can cause problems with usability and reach of target users. It is common for a research group to develop their own analysis pipeline or a method to tackle an issue that is specific to their research question, especially if the team happens to have an in-house bioinformatician. After the project is done, it is common for the tool to be published with it, often in the form of a pipeline or a command line tool that is uploaded to an external code repository. These sorts of tools often do not reach the potential user base as the announcement of it is simply buried within a publication of the main research. Additionally, time constraints usually do not allow for the developer of the tool to put large amounts of time to developing the user interface and general usability of the tool. Rather, the software is often supplied "as is". These types of programmes and tools are being released at an incredible pace, and a lot of them will get lost in the flood. At best, teams can have the resources to develop and release a web service for their tool, which often results in a better user experience. However, these services can be troublesome to upkeep for longer if they are self-hosted, and often times they are simply taken out of service.

Arguably, some of the best tools for bioinformatics come from organizations that have the resources to spend real development time on the service. These sorts of tools are often commercial, such as the standalone software package Geneious Prime, or backed by an organization that is able to fund the development of the web service, as with the web service Galaxy (Table 1). Some less resource rich tools can survive in the wild, although they are often backed up by a considerable community, as seen for example with BEAST. It can of course be the case that a lone researcher or a smaller team develops a brilliant bioinformatic tool, even if the usability might be lacking. Luckily bioinformatic platforms such as Geneious and Galaxy have some capability of integrating popular tools to their services, which can help expand the visibility and lifetime of the lesser seen tools. This solution does not necessarily fix the usability problem though, and also has the risk of hiding the tool in question behind a commercial licence. Furthermore, web services such as NextStrain and Galaxy can be excellent tools for specific tasks. Having stated the above, web services can also suffer from outages and performance issues, and sometimes they simply cannot carry out the more computationally or data heavy tasks.

3 Aims

The publications included in this dissertation were bundled into three groups, each with a distinct theme. The specific research aims for each group were:

1. To perform comparative *in vitro* analysis of melanoma cell-adapted echovirus 7 (E7) virotherapy drug, named Rigvir®, against clinical E7 isolates, and carry out sequencing and bioinformatic analyses of Rigvir® to investigate potential molecular and structural mechanisms that might explain Rigvir®'s claimed oncolytic effects (**I** & **IV**).
2. To perform analysis of coxsackievirus A9 (CVA9) RGD receptor binding site conservation and variability. Establish phylogenetics of the CVA9 data set against publicly available CVA9 data and investigate recombination effects through topological incongruence within phylogenies. Additionally, the aim was to analyze available CVA9 clinical symptom data to investigate and establish links between phylogenetic information and observed symptoms (**II**).
3. To develop broadly reactive monoclonal antibodies against parechovirus (HPeV) VP0 (**III**). Additionally, to isolate and perform structural analysis of the binding site for phage display-derived single-chain variable fragment (scFv) antibodies against HPeV-1 VP0 (**V**).

4 Materials and Methods

As publications included in this thesis included multiple different picornaviruses and types of studies, the methods described below have been bundled based on their respective type. The publications to which each chapter or paragraph refers to have been marked with roman numerals **I**–**V**.

4.1 Viruses, virus cultures, antibodies, and proteins (I-V)

In **Publication I**, vRNA was extracted from infected rhabdomyosarcoma (RD) cell lysates using the E.Z.N.A. viral RNA kit (Omega Bio-tek, Norcross, GA, USA). Likewise, vRNA of Rigvir® was extracted directly from the drug virus ampoule with the same kit.

In **Publication II**, vRNA from CVA9 samples was extracted using a QIAamp MinElute Virus Spin Kit (Qiagen GmbH, Düsseldorf, Germany) as instructed by the manufacturer.

In **Publications III** and **V**, HPeV-A1 (Harris) and HPeV-A2 (Williamson) prototype strains were purchased from ATCC. HPeV-A1 (152478, 452252 and 350757), HPeV-A3 (152037 and K251181-02), HPeV-A4 (K251176) and HPeV-A5 (20552322) isolates were gifts from Dr. Katja Wolthers (AMC, Amsterdam, the Netherlands). HPeV-A1 (153-20, 101-17, 103-2, 125-7, 145-12, and 150-8) were from National Institute of Health and Welfare (Helsinki, Finland). HPeV-A1 isolates FI0003, FI0007, FI0008, FI0111, FI0114, FI0219, FI0222, FI0433, FI0435, FI0578 and HPeV-A6 types (FI0147 and FI0189) were from the laboratory collections located at the University of Turku (Kolehmainen et al., 2012). Enteroviruses (CVA9, CVB2, CVB3, CVB5, E2, E11, E20 and E30) and rhinoviruses (RV-A1b, RV-A10, RV-A24, RV-B14, RV-B35 and RV-B86) were from ATCC. In **Publication III**, the virus inoculation experiments were conducted in human epithelial lung carcinoma (A549), human cervical cancer (HeLa), colorectal adenocarcinoma (HT-29), and RD cell lines, purchased from ATCC, and maintained in Dulbecco's Modified Eagle medium (DMEM) supplemented with 10 μg ml^{-1} gentamicin, 2 mM L-glutamine, and 10% Fetal Bovine Serum (FBS), incubated at 37 °C (5% CO_2). In **Publication**

V, virus inoculation experiments were conducted in HT-29 and RD cell lines and maintained in DMEM as above.

In **Publication IV**, viruses used in the study included E7 isolated from the original Rigvir® ampoule (Sia Latima Ltd., Riga, Latvia), a prototype "Wallace" E7 isolate was purchased from ATCC, and four clinical E7 isolates originating from Finland were included from the University of Turku laboratory collection (98-57213, 98-59065, 98-60628, and 07VI447). All viruses were typed as "echovirus 7" according to the typing criteria set by the ICTV (Zell et al., 2017). The viruses were inoculated onto RD cells, and cell lysate stock viruses were collected three days post-infection based on cytopathic effect (CPE). Human foreskin fibroblasts (HFF) were maintained in M199 medium supplemented with 5% FBS and 10 µg ml^{-1} gentamicin. Human bronchial epithelial cell line (16HBE14o) was maintained in Minimum Essential Media (MEM) with 10% FBS, 2 mM L-glutamine and 10 µg ml^{-1} gentamicin (Gruenert et al., 1988). Human cervical cancer (HeLa) cell line was maintained in Basal Medium Eagle (BME) with 7% FBS and 10 µg ml^{-1} gentamicin. Human epithelial lung carcinoma (A549) cell lines were maintained in Ham's F12 medium with 7% FBS and gentamicin. RD, human glioma (U373MG), human hepatocarcinoma (Huh7), human colorectal adenocarcinoma (SW480) and human breast cancer (MCF-7) cell lines were maintained in DMEM supplemented with 10% Fetal Calf Serum (FCS) and 10 µg ml^{-1} gentamicin. Viral RNA was extracted from 150 µL volume of virus-infected RD cell lysates using E.Z.N.A vRNA kit (Omega Bio-tek, Norcross, GA, USA) according to the manufacturer's protocol, and stored frozen at -80 °C.

In **Publication V**, anti-FLAG tag antibody (used for the FLAG tag on scFv antibodies) was obtained from Cell Signaling Technology (#2368). Light Diagnostics pan-enterovirus reagent (#3360) was obtained from Millipore. Horse radish peroxidase (HRP)-conjugated anti-rabbit antibody was obtained from Jackson Immuno Research (111-035-003), and TMB substrate was obtained from Thermo Fisher (34018). Alexa Fluor 488-labeled anti-rabbit and anti-mouse secondary antibodies (Life Technologies #A-11008 and #A-11001, respectively), as well as DAPI (Sigma Aldrich D1306) were used in immunofluorescence assays. Finally, goat anti-rabbit secondary antibody (Licor IRDye® 800CW) was used in Western blotting.

4.2 Infection assays (IV)

Cells were seeded at 10,000 per well and grown on 96-well plates (PerkinElmer Health Sciences, Turku, Finland) to 60%–80% confluency. Viruses were titrated at 1:10 intervals to reach optimal infectious dose. This was carried out to avoid cytotoxicity effects of cell lysate virus stock. Cells were visually inspected for signs

of CPE for 3 days post-infection. Subjective terminology with (−; no cytolysis), (−/+; borderline result), (+; few cells lysed), (++; 50% of cells lysed) and (+++; full cytolysis of cells) were used to determine the extent of CPE.

4.3 Polymerase chain reaction and real time quantitative RT-PCR (II, IV)

In **Publication II**, previously described primers (McWilliam Leitch et al., 2009) for VP1 and 3Dpol and modified semi-nested PCR protocol were used. Briefly, reverse transcription (RT) step was performed in 10 µL volume and contained 0.5 µL ImProm II (Promega, Madison, WI, USA), 1 µL outer antisense (OAS) primer at 1 µM final concentration, ImProm II buffer, RNasin, and ddH_2O. Samples were incubated at 42 °C for 60 min. Five microliters of produced cDNA was added to 25 µL PCR1 reaction, which included 0.5 µL Platinum Taq High Fidelity enzyme, 2.5 µL primers at 1 µM final concentration, buffer, and ddH_2O. Cycling conditions for both VP1 and 3Dpol amplification were 95 °C for 3 min, then 35 cycles of denaturation at 95 °C for 20 s, annealing at 50 °C for 30 s and extension at 72 °C for 60 s. Final extension stage was for 5 min. In snPCR2, 1 µL of PCR1 reaction product was used.

In **Publication IV**, the specimens were analyzed using a real-time RT-PCR (Peltola et al., 2008). RT-PCR amplification was performed with conserved picornavirus primers from the 5′ noncoding region of the genome (Lönnrot et al., 1999). PCR reactions were performed in a RotorGene 6000 instrument (Qiagen, Hilden, Germany) in 25 µL reactions containing 5 µL of the cDNA, QuantiTect SYBR Green PCR mix (Qiagen, Hilden, Germany), and 600 nmol/L of virus primers. The PCR protocol used consisted of the following steps: 95 °C for 15 min, followed by 45 cycles at 95 °C for 15 s, 65–55 °C for 30 s (touchdown 1 °C /cycle for the first 10 cycles), and 72 °C for 40 s (melt 72–95 °C, 0.5 °C /s). Relative cycle threshold (Ct) values were determined from the results for each sample. The values were used to standardize the amount of virus inoculated onto RD cells. The cells were lysed and RNAs were extracted at 1 h and 72 h time points and Ct values were determined as shown above.

4.4 Monoclonal antibody production (III)

Recombinant GST-HPeV-A1-VP0 fusion protein and GST protein used as a control were purified as previously described (Alho et al., 2003) with minor modifications. Briefly, cell pellets were suspended and lysed in Bugbuster buffer (#70584-4, Novagen, USA) for 30 min on ice. Bugbuster buffer was supplemented with 2 mg/mL lysozyme (#L6876, Sigma, USA), Benzonase (#B16012, Biotool, USA) and

protease inhibi- tor (#A32963, Thermo-Fischer, USA). The supernatant was cleared by centrifugation at 10000 rpm for 10 min at 4 °C. Supernatant was collected and incubated with glutathione sepharose (#GE17-0756-01, GE Healthcare, USA) in a column for 2 h at room temperature. After extensive washing, column elution was done with 20 mM Glutathione (#6200202, Bio-Rad, USA) prepared in PBS buffer, pH 8.0. The eluates were collected and buffer was exchanged to PBS by dialysis overnight at 4 °C. The purity of eluted protein fractions was analyzed with SDS-PAGE.

Two Balb/c mice were immunized with 20 µg of HPeV-A1-VP0 recombinant protein in FCA adjuvant. Four booster injections were given in two weeks' intervals. Mice were sacrificed two days after final injection and splenocytes were fused to mouse myeloma cells using ClonaCell HY Kit (Stemcell Technologies, Canada). One thousand (1000) cell clones were cultivated on 96-well plates and tested for HPeV-A1-VP0 on ELISA and immunofluorescence assay (IFA) in HPeV-A1 (Harris strain) infected HT-29 cells. Twelve (12) clones were selected for further assaying of which one IgG and one IgM clone specific for HPeV-A1 was isolated. IgG clone was chosen for further characterization. Antibody from this clone was isotyped as IgG1 κ and named Mab-PAR-1. Mab-PAR-1 was purified in CIM r-Protein G column (BIA Separations, Slovenia) using ÄKTA FPLC (GE Healthcare, USA). Protein purity was determined using SDS-PAGE and concentration was determined by BCA assay.

4.5 Phage-display derived scFv antibody production (V)

HPeV-A1-VP0 was biotinylated with Ez-link sulfo-NHS-biotin (Thermo Fisher Scientific) according to manufacturer's instructions. HPeV-A1-VP0 specific binders were selected by phage display from synthetic single-chain variable fragment (scFv) phage libraries ScFvM16 and ScFvP33. Both libraries use a single human scFv gene as a framework and display the antibody as a fusion to the truncated form of M13 phage coat protein p3. The libraries contain diversity at the hypervariable CDR regions, but have distinct binding profiles (Huovinen et al., 2013). The methods used for M13 phage display have been previously described (Huovinen et al., 2013). In brief, purified, biotinylated recombinant HPeV-A1-VP0 was first immobilized on Dynabeads® M-280 paramagnetic streptavidin beads (Life Technologies). Before the actual panning step where the phage antibody library was incubated with bio-GST-VP0 bound to the M280-beads, negative selections against mere M280-beads and the beads with bio-GST were done to deplete the library from binders against streptavidin and GST. In the panning, 10 µM biotin and 50 µg/ml GST were added as blockers to prevent bead clustering through the biotinylated antigen and

enrichment on binders against GST in the panning antigen. At the first selection round 5×10^{12} colony-forming units (cfu) of both scFv phage libraries were used as mixed. The mass of antigen-coupled beads used for the first and second round selections was 0.5 mg or 0.05 mg, respectively. The phages were incubated with the beads in TBS (50 mM Tris, 150 mM NaCl, 1% BSA, pH 7.5) containing 1% bovine serum albumin (BSA) for 30–60 min at room temperature with rotation. The unbound phages were removed by washing three times with the same buffer, followed by one wash with TBS + 0.05% Tween-20. Elution of the bound phages was performed with trypsin. Enrichment of specific phages was monitored by a phage immunoassay.

The gene cassette in the vector pLK06FT has the following orientation: SfiI (restriction enzyme site)—scFv—SfiI/AvaI—AP (bacterial alkaline phosphatase)—TEVPro (protease cleavage site)—His6—Gly-Ser-Gly-linker—FLAG—STOP—HindIII.

After two and three rounds of phage display selection, the scFvs genes originating from the human synthetic scFv antibody library were cut off from the phagemid vector (pEB32x/scFv) by SfiI digestion and cloned as a pool into the vector pLK06FT for single-clone immunoactivity screening16. Electro-competent *E. coli* XL1-Blue cells (Stratagene) were transformed, and scFv-AP fusion proteins with C-terminal His6 and FLAG tags were expressed in the volume of 200 µl in a 96-well format as described earlier36. The immunoreactivities of scFv antibody fragments were determined on streptavidin micro- titration plates (Uniogen Oy, Turku, Finland) bound with biotinylated GST-VP0. As a control, binding was also measured to biotinylated GST and streptavidin well only. ScFv-APs were bound from lysed culture supernatants for 1 h. After four washes pNPP substrate was added, plates were shaken 3 h at room temperature and A405 was measured. In second screening step binding of eleven IMAC (immobilized metal affinity chromatography) -purified scFv-antibody clones to VP0 protein was confirmed using Eu-labelled anti-AP antibody in a time-resolved fluorescence (TRF) assay. The scFv-HPeV antibodies were expressed in E. coli periplasm in 50 ml shake flask cultures induced o/n at 26 °C with 100 µM IPTG. Cells were harvested and lysed, and scFv was IMAC-purified with HisPur™ Ni-NTA Spin Columns, 0.2 mL resin bed (Thermo Scientific, Espoo, Finland).

4.6 ELISA (III, V)

In **Publication III**, wells of 96-microtiter plate (#3590, Corning, USA) were coated with Mab-PAR-1 (500 ng/well) in bicarbonate buffer (100 mM, pH 9.6) and incubated overnight at 4 °C. Next day, the wells were washed with PBS-T buffer (PBS supplemented with 0.05% Tween 20, pH 7.4) and blocked with 3% BSA at

room temperature for two hours. After three washes with PBS-T buffer, HPeV-A1- or mock-inoculated cell lysates were pipetted onto wells coated with Mab-PAR-1. Purified VP0-GST and GST proteins were used as positive and negative controls, respectively. After one-hour incubation at room temperature, wells were washed four times with PBS-T and rabbit anti-HPeV-A1 polyclonal serum was added to the wells. After one hour of incubation, wells were washed four times with PBS-T and HRP-conjugated Goat anti-rabbit IgG antibodies (#111-035- 003, Jackson Immunoresearch, USA) were added and kept at room temperature for one hour. Wells were washed three times with PBS-T buffer, TMB substrate (#34028, Thermo-Fisher, USA) was added, and reaction was stopped with addition of 0.45 M sulphuric acid. OD450 values were measured with Victor^3 multilabel counter (PerkinElmer, USA).

In **Publication V**, Corning High binding 96 well plate (#3590) was coated with three scFv-HPeV antibodies (scFv-55, -59, and -71) (500 ng/well) in PBS and incubated at 4 °C overnight. The plate was washed three times with PBS and blocked with 3% BSA in PBS for two hours at room temperature followed by washing with PBS. Purified HPeV-A1 (Harris strain), control cell lysate, purified VP0-GST fusion protein and purified GST were diluted with PBS and added onto the wells (200 ng/well in PBS). After 90 min incubation at room temperature, the plate was washed three times with PBS, and rabbit anti-HPeV-A1 antiserum (pAb-HPeV-A1) was added onto the wells. Antiserum was incubated 1 h at room temperature followed by washing with PBS. HRP-conjugated anti-rabbit secondary antibody was added and incubated for 1 h followed by washes with PBS. TMB substrate (Pierce, USA) was added onto the wells and incubated at 37 °C for 30 min which after 0.45 M H2SO4 was added. OD450 was read with Victor^3 multilabel counter (PerkinElmer, Turku, Finland).

4.7 Western blot (III, V)

In **Publication III**, HPeV-A1 infected or non-infected A549 cell-lysate, purified VP0-GST fusion protein and purified GST protein were prepared individually with SDS sample buffer and run in 4–20% SDS PAGE (4561093S, Bio- Rad, USA). Page ruler pre-stained (#PI26619, Thermo Scientific, USA) was used as a protein size ladder. Separated proteins were transferred to nitrocellulose membrane (#10600001, Life Science, USA). The blotted membranes were blocked overnight with TBS containing 5% BSA followed by blotting with 10 µg/mL Mab-PAR-1 in 1% BSA in TBST for 1 h. After washes with TBST, membrane was incubated with anti-mouse IRDye® 680RD (#926-68070, LI-COR Biosciences, USA) for 1 h followed by washes and analyzing with Odyssey (LI-COR, USA).

In **Publication V**, lysate of HPeV-1 infected HT-29 cells, purified VP0-GST fusion protein, HT-29 cell lysate and purified GST protein were heated in SDS sample buffer at 95 °C for 5 min, and run in 4–20% SDS PAGE (Bio-Rad 4–20% Mini-Protean® TGX™). Chameleon Duo (Licor 928-60000) was used as a marker. Separated proteins were transferred to membranes (Amersham™ Protran™ Nitrocellulose Blotting Membrane, Life Science #10600001) with TransBlot® Semi-Dry Transfer Cell (Bio-Rad). The blotted membranes were blocked overnight with TBS containing 5% BSA followed by blotting with 10 µg/ml of either scFv-55, -59 and -71 antibody in 1% BSA in TBS-T for 1 h. After washes with TBS-T, membranes were incubated with anti-FLAG antibody for 1 h. Membranes were then incubated with goat anti-rabbit secondary antibody (Licor IRDye® 800CW) for 1 h followed by washes and analysing with Odyssey (Licor).

4.8 Immunofluorescence microscopy (III-V)

In **Publication III**, A549 or HT-29 cells seeded onto Viewplate-96 Black Clear Bottom plates (#6005182, PerkinElmer, USA) were infected with different dilutions of HPeV-A isolates or types aiming to achieve 20–40% infection per cell count in the well. The infection was allowed to proceed for six hours at 37 °C in incubator and 5% CO_2. After six hours of infection, the cells were washed once with sterile PBS. Fixing of cells was carried out using 4% formalin and permeabilization with 0.2% Triton-X100. Staining with Mab-PAR-1 antibody (1 µg/mL), Pan-Enterovirus Re- agent (#3365, Light Diagnostics, USA) or mouse pan-entero monoclonal antibody (#C01700 M, Meridian, USA) in 3% BSA in PBS at room temperature for one hour. After incubation and washing, cells were stained with Alexa Fluor 488-labelled anti-rabbit (#A-11008, life- technologies, USA) or anti-mouse (#A-11001, Life-technologies, USA) secondary antibody for one hour at room temperature. The nuclei were stained with DAPI (#D1306, Sigma, USA). EVOS FL-Auto was used to visualize the cells.

In **Publication IV**, cells were seeded at 10,000 per well and grown on 96-well plates (PerkinElmer Health Sciences, Turku, Finland) to 60% confluency. Viruses were inoculated onto cells, and unbound virus was removed after 1 hour (h) of incubation by washing three times with medium. Virus inoculum was defined as the volume that resulted in approximately 50% cell infectivity (Median Tissue Culture Infectious Dose; TCID50) in RD cells based on the titration assay. Fresh medium was added, and the cells were incubated at 37 °C. The infection was allowed to proceed for 6 h, after which the cells were fixed with 4% formalin and permeabilized with 0.2% Triton X-100. Infected cells were stained with monoclonal pan-enterovirus 9D5 antibody (Millipore, Burlington, MA, USA) followed by combined staining with Alexa Fluor 488-labeled secondary anti- mouse antibody and staining

of the nuclei with DAPI (25 µg ml−1). Images were acquired using a Zeiss Axiovert 200Mmicroscope equipped with A-Plan 10×/0.25 Ph1Var1 objective (Zeiss, Oberkochen, Germany). Brightness and contrast levels of the images were adjusted with Fiji (ImageJ) (Schindelin et al., 2012) and Adobe Photoshop image analysis programs (Adobe Systems Inc., San Jose, CA, USA). Likewise, quantitative image analysis of the immunofluorescence images was carried out in Fiji (ImageJ) by first performing noise and background removal on the individual image channels. Segmentation of the nuclei was performed automatically, and the results were manually checked for consistency. After watershedding, the particle analysis function in Fiji was then used together with appropriate size constraints to calculate the number of individual nuclei. The number of infected cells was determined as largely similar by segmentation and particle analysis. However, the segmentation was performed more manually in cases where the fluorescence channel signal did not result in a good segmentation result otherwise. In cases where a large portion of the cells were infected and manual segmentation was infeasible, Cellpose (Stringer et al., 2021) was used to automatically segment the infected cells and to create a mask of the area. The resulting image mask was then manually checked and adjusted if needed. Afterwards, the mask of the infected cells and the segmented nuclei image were combined in the Fiji image calculator through the "AND" function. The image calculator resulted in an image showing only the areas where the infected cell mask and the nuclei segmentation images overlapped, resulting in a new segmentation of cell nuclei belonging only to the infected cells. The nuclei were then calculated as described before using the Fiji particle analysis function.

In **Publication V**, DMEM was supplemented with 1% FBS. Cells were seeded at 10,000 per well and grown on black 96-well plates (6005182, Perkin Elmer) to 90% confluence. The cells were infected with a virus dilution aiming at infection occurring in 20–40% of cells in order to facilitate better separation of fluorescence-forming units and subsequent image analysis. The infection was allowed to proceed for 6 h at 37 °C and 5% CO_2 followed by fixing with 4% formalin and permeabilization with 0.2% Triton X-100. Infected, fixed and permeabilized cells in 96-well plates were stained as follows: scFv-55, -59 and -71 antibodies were diluted to 3% BSA in PBS to concentration of 1 µg/ ml and incubated on the cells for 1 h. After washes with PBS, anti-FLAG antibody was added and incubated for 1 h followed by washes with PBS. Alexa Fluor 488-labeled anti-rabbit secondary antibody was incubated for 1 h. With enterovirus testing pan-Enterovirus antibody was used as a primary antibody and Alexa Fluor 488-labeled anti-mouse was used as a secondary antibody. The nuclei were visualized with DAPI. 96-well plates were visualized by fluorescence microscope using EVOS FL Auto with 10 × objective (Thermo Fisher Scientific), and the brightness and contrast were adjusted with EVOS FL and Gimp 2.0. software.

4.9 Epitope mapping (III, V)

In **Publications III** and **V**, epitope mapping was performed at Pepscan Presto BV (Lelystad, The Netherlands) using Pepscan's proprietary Chemically Linked Peptides on Scaffolds (CLIPS) technology. 15-mer peptides from the VP0 protein (corresponding to nucleotides 710–1576 of HPeV-A1 Harris strain (GenBank acc no L02971)) were generated with an offset of one amino acid. Mab-PAR-1 antibody was incubated with the peptides immobilized on a glass slide (peptide library) in 4% horse serum, 5% ovalbumin (w/v) in PBS/1% Tween. After washing, the peptide library was incubated with a 1/1000 dilution of rabbit anti-mouse IgG (H + L) HRP conjugate for one hour at 25 °C, washed again, and incubated with the peroxidase substrate 2,2'-azino-di-3-ethylbenzthiazoline sulfonate (ABTS) and 2 µl of 3% H_2O_2. After one hour, the color development was measured and quantified with a charge coupled device (CCD) camera and an image processing system (Slootstra et al., 1996).

4.10 Sequencing (I-II)

In **Publication I**, viral RNA was prepared for sequencing using the NEBNext Ultra RNA library preparation kit (catalog number E7530, New England BioLabs) and NEBNext multiplex oligos for Illumina adapter kit (catalog number E7335) and sequenced on the Illumina MiSeq platform. Contigs were assembled using *de novo* assembly protocols (MIRA version 4.0.2).

In **Publication II**, amplified DNA from RT-PCR reaction was treated with ExoSAP mixture and directly sequenced using 3730xl DNA Analyzer (Eurofins GATC, Köln, Germany) with the inner sense or antisense primer used for amplification (Eurofins Genomics, Ebersberg, Germany). Overlapping sequences were assembled with Staden Package, trimmed for equal length and aligned with Clustal omega implemented in Seaview program (Gouy et al., 2010), and subjected to evolutionary analyses. All newly generated sequences were deposited in the GenBank database under the accession numbers MN493979-MN494086.

4.11 Sequence analysis (I-II, IV-V)

In **Publication I**, sequence and phylogenetic analyses were conducted using BioEdit v7.2.5 (Hall, 1999) and MEGA v7) (Kumar et al., 2016).

In **Publication II**, to extract the RGD motif from the sample sequences, the raw inner antisense (IAS) and inner sense (IS) files were loaded into Unipro UGENE program (Okonechnikov et al., 2012) where VP1 open reading frames (ORFs) containing the motif were extracted if the sequence data were found to cover the area adequately. These sequences were later translated and trimmed to a 20 aa window

around the motif for further analysis with the BioEdit v7.2.5 software (Hall, 1999). Possible unresolved bases found in some sequences, likely due to RGD motif's position in the C-terminal end and the Sanger sequencing method used, were manually fixed by using the data from the original sequencing chromatograms. Eventually, a total of 64 sequences were used in the final RGD motif analysis, including the prototype Griggs (GenBank accession number D00627.1) sequence, as well as all available full length CVA9 sequences from GenBank. The putative HS binding site region was extracted from the originally aligned sequence data with a 29 aa window around the VP1-T132 position as suggested in previous studies (McLeish et al., 2012; Merilahti et al., 2016). The resulting alignment was run through protein BLAST using the non-redundant protein sequence database, excluding CVA9 sequences, with the aim of identifying other potential enteroviruses with present HS binding sites. From the top 250 returned hits, the non-redundant enterovirus sequences were combined with the putative HS region from the full length CVA9 GenBank sequences included in this study and subjected for further analysis.

In **Publication IV**, viral genome sequences from **Publication I** (GenBank acc. no. MH043132-MH043137) were analyzed further. Sequence and phylogenetic analyses were carried out using Geneious Prime 2022.0.1 (https://www.geneious.com, accessed on 19 February 2022). Sequence alignments within Geneious Prime were conducted using the MUSCLE (Edgar, 2004) and MAFFT (Katoh & Standley, 2013) plugins for nucleotide and amino acid alignments, respectively, with the suggested settings. Phylogenies were built using Geneious Tree Builder with bootstrapping, using 500 replicates and a support threshold of 70%. Additionally, echovirus 9 prototype sequence X92886 (strain Barty) was set as an outgroup of the phylogeny.

In **Publication V**, all non-redundant HPeV sequences covering the VP0 region were retrieved from GenBank (n=361) and combined with a sequenced set of parechoviruses of types 1 and 3–6 (n=24) from this study. Multiple sequence alignments were carried out with Geneious Prime 2021.0.325 using the MUSCLE (Edgar, 2004) plugin with default settings.

4.12 Structural analysis (III-V)

In **Publications III** and **V**, structural analyses of the epitope were carried out using UCSF Chimera v1.15 (Pettersen et al., 2004). The structure of HPeV-1 virion obtained from the Protein Data Bank (PDB ID: 4Z92) (Berman, 2000; Kalynych et al., 2016) was used as a reference to determine the exact location and conformation of the epitope.

In **Publication IV**, structural analysis of the DAF region was carried out using UCSF ChimeraX v1.2 (Pettersen et al., 2021) using the crystal structure of E7 capsid proteins VP1-VP4 published previously by Plevka et al. (PDB ID: 2X5I) (Plevka et al., 2010). Amino acid mutations unique to Rigvir®, as discovered through previous sequence analysis, were mapped onto the crystal structure against residues that were previously described to form the DAF–E7 contact interface. Additionally, the surface exposure of the mutations was analyzed to determine their potential ability to affect receptor binding through direct surface exposure.

4.13 Phylogenetics (II, IV)

In **Publication II**, bootstrapped phylogenetic trees for the VP1 and 3Dpol regions were constructed using the MEGA6 software package. Maximum likelihood (ML) and maximum composite likelihood (MCL) methods were used in the phylogenetic tree construction using the Kimura 2-parameter substitution model (Kimura, 1980) with 500 bootstrap replicates. Data used in the analysis included all codon positions, with pairwise deletion for missing data. A Markov Chain Monte Carlo (MCMC) method, implemented in the BEAST software package version 1.8.2 (Drummond et al., 2012), was used for temporal phylogenetic analysis of the data. For the VP1 region, dated sequences were analyzed with a chain length of 60 million, sampling every 6000 states, under the SRD06 substitution model (Shapiro et al., 2006) with the assumptions of a relaxed molecular clock model and a constant size coalescent. Other parameters were initially manually checked and adjusted, with final optimization done during the burn-in period. For the 3Dpol region, BEAST analysis was run with a chain length of 90 million to ensure good quality control parameters. Otherwise, the process for the 3Dpol region analysis was carried out the same way as for the VP1 region. Outputs from BEAST for both regions were analyzed within the TRACER v1.6 program to ensure convergence through graphical checks, as well as adequate quality control parameters of posterior distribution (ESS > 200). Final phylogenetic trees from BEAST outputs were constructed with the TreeAnnotator v1.8.2 software to find the maximum clade credibility tree from all the sampled tree states. Finally, the obtained trees were visualized with the FigTree v1.4.2 software.

In **Publication IV**, a neighbour-joining tree was built using 500 bootstrap replicates with Geneious Prime and the Geneious Tree Builder (*Geneious Prime 2021.0.3*, n.d.), using all of the non-redundant full-length E7 sequence data available from GenBank (n=30).

5 Results

5.1 Analysis of the oncolytic virotherapy drug Rigvir® against clinical echovirus 7 isolates

5.1.1 Sequence analysis and phylogenetics of Rigvir and other echovirus 7 isolates

Rigvir® is virotherapy drug that is derived from a clinical echovirus (E7) isolate. Specifically, it has been cell-adapted in melanoma cells through multiple passaging rounds, and is being marketed as an effective oncolytic virotherapy especially against melanoma (Alberts et al., 2018; Tilgase et al., 2018).

The sequence data obtained from Rigvir® and all other full-length E7 sequence data (n=30) was subjected to sequence and phylogenetic analyses using the MUSCLE (Edgar, 2004) and MAFFT (Katoh & Standley, 2013) software for alignment, and Geneious Prime (*Geneious Prime 2021.0.3*, n.d.) software for general bioinformatics including sequence analysis and phylogenetics. In respect to the Wallace prototype sequence, the sequence analysis revealed that the whole genome nucleotide sequence identity ranged between 79.1% to 99.9%, while within the open ORF the nucleotide identity ranged between 78.9% to 99.8%. The amino acid identity within the ORFs ranged from 95.8% to 98.5%. Additionally, similar high conservation of amino acid identities was observed within only the capsid proteins VP1 – VP4, with identities ranging from 96% to 97.1%. Results showed that the most variable region within the capsid proteins was located in the VP1 protein encoding region, especially at the 5´ and 3´ ends. The amino acid sequence identity within the non-structural proteins ranged from 92.6% to 98.7%, with the most variable region being the 2A protein encoding region. An analysis of mutations present in Rigvir® was focused on the capsid proteins VP1 – VP4, as they have the highest potential of directly affecting tropism. Rigvir® was observed having a total of nine mutations in these proteins (Table 2). Two of the mutations were in VP1, five in VP2, and two in VP3. Sequence analysis of Rigvir®'s sequence did not exhibit any other major differing features that were noteworthy within the examined genome areas.

Table 2. Unique amino acid mutations present in Rigvir® capsid proteins. Sequence positions indicated as in echovirus 7 crystal structure (PDB ID: 2X5I).

CAPSID PROTEIN	ECHOVIRUS 7 RESIDUE	STRUCTURE POSITION	RIGVIR® RESIDUE	SURFACE EXPOSURE
1	I/L	76	F	Yes
1	V	153	I	Yes
2	L	17	H	No
2	K/N/L	138	R	Yes
2	E/Q/N	162	G	Yes
2	V	176	I	No
2	T/A	231	N	Yes
3	I	158	L	No
3	A	235	T	Yes

5.1.2 Structural analysis of Rigvir

Structural analysis of Rigvir® was carried out using UCSF ChimeraX v1.2 (Pettersen et al., 2021) in order to determine the exact structural location of the mutations observed in Rigvir®'s capsid proteins, as well as to evaluate their potential effects based on currently available information. Structural analysis of the mutations was performed on the resolved crystal structure of an E7 (PDB ID: 2X5I). Out of the nine mutations found previously, six were exposed on the surface on the capsid (**IV**; Fig. 3B). One of these mutations, E/Q/N162G in VP2, was positioned in a location where previous studies had mapped residues that form the contact interface with the decay-accelerating factor (DAF) receptor (Plevka et al., 2010). The mutation was spatially close to another mutation, K/L/N138R, which was positioned in the general structural region involved with DAF contact, however, the specific residue was not previously mapped to be involved in the direct interaction (Plevka et al., 2010).

5.1.3 Cellular assays

Cellular assays were performed to compare Rigvir®'s *in vitro* performance against the four clinical E7 isolates, as well as the prototype Wallace isolate, previously sequenced in **Publication I**. Eight cell lines in total were infected, with two native non-cancer primary cell lines (HFF and 16HBE14o) and six cancer cell lines (HeLA, A549, SW480, Huh7, U373MG, and MCF-7).

The viruses were first titrated in RD cells to establish a baseline TCID50 for all isolates, after which comparable levels of virus was used to infect each cell line to determine the viability in each line based on observed CPE three days post-infection.

While observed CPE between the various cell lines was expectedly different, no major differences were observed between the viruses within each cell line. Interestingly, in the breast cancer cell line MCF-7 did not show CPE with any isolate. Furthermore, both of the native non-cancer cell lines were infected by all of the isolates.

To confirm the previous results and to ensure that infection was not only taking place in MCF-7 without cytolysis, and to visualize the infection within the interior, IFA was used to observe the infections using the TCID50 established earlier. The IFA results were in line with the observed CPE results, confirming that MCF-7 was in fact not being infected by any of the isolates. Importantly, the microscopy results also confirmed that Rigvir® indeed infects also native non-cancer cell lines in a similar fashion as did the clinical E7 isolates. The immunofluorescence microscopy also enabled more accurate image analysis to calculate the infectivity of each virus as the percentage of total cells infected. Based on the image analysis, Rigvir® was seen infecting Huh7 cells slightly more efficiently than other viruses, with 48.9% of cells infected, while the second most infective virus was the clinical isolate 07VI447 at 40.6%. In other cell lines, Rigvir® either performed similarly or worse than other clinical isolates.

Lastly, as Rigvir® is a cell-adapted E7, it was speculated that the it would be possible for Rigvir®'s cytolytic properties to be more pronounced than those of the clinical isolates, which would in turn enhance its ability to infect and destroy the cells susceptible to infection. To examine possible differences in relative quantities of virus multiplication, RT-qPCR was performed with each virus, isolated from RD cell lysate. The samples were collected at 1 h and 72 h time points post-infection, and the relative Ct value from the RT-qPCR assay was used to calculate increases in virus amounts. Rigvir® could not be seen deviating from the clinical isolates also in this case, and the observed virus amounts were in line the previous results as compared to the other viruses.

5.2 Coxsackievirus temporal phylogenetics and receptor binding site analyses

5.2.1 Coxsackievirus A9 phylogenetics

CVA9 is a particularly pathogenic enterovirus type compared to other CVAs, with previous studies linking it more closely to the more pathogenic group B coxsackieviruses (Aoki et al., 2012; Cui et al., 2010; Palacios & Oberste, 2005). As such, it is an interesting target for bioinformatic analyses.

A unique data set of 75 CVA9 sequences collected between 1959 and 2016 was subjected to recombination and sequence analyses. The approach to detecting

recombination events in this case was focused on utilizing the VP1 and 3D genomic regions of the viruses, and looking for changes in the topologies of the trees resulting from phylogenetic analysis. At the same time, the phylogenetic analysis carried out provided estimates of the most recent common ancestor (MRCA) of the isolates, and enabled the examination of possible geographical relationships through a longer time span. Additionally, the addition of patient symptom data allowed for examining if pathogenicity can be linked to phylogenetics in the case of CVA9. The data set consisted of 48 isolates from Europe, 12 from Americas, and 15 from Asia-Pacific, including the prototype CVA9 sequence "Griggs".

Temporal phylogenetic analysis of the data set with the BEAST software package v1.8.2 (Drummond et al., 2012) was able to estimate the MRCA, as well as the evolutionary rate of the viruses. The VP1 data estimated the MRCA to approximately 1889 (95% CI 1859–1918), with a mean substitution rate of 4.1×10^{-3} substitutions/site/year (95% HPD range $3.1–5.0 \times 10^{-3}$). For the 3D region, the MRCA was estimated further back to 1814 (95% CI 1726–1886), with a mean substitution rate of 3.4×10^{-3} substitutions/site/year (95% HPD range $2.3–4.6 \times 10^{-3}$).

The phylogenetic trees obtained from BEAST showed clear clustering based on the country of origin, and to some degree, on the sample collection year. Overall, the VP1 region was divided into six different major clades based on the tree topology (**II**; Fig. 1A). Corresponding clades were identified from the 3D phylogeny, and incongruence of previously established VP1 branch relationships in the 3D phylogeny was examined as evidence of recombination. Interestingly, examining the VP1 phylogeny by itself, the more recent isolates collected from the Americas between 2005 and 2016 did not group with the older American isolates. Rather, most of these more recent isolates grouped more together with sequences from Asia-Pacific and Finland in clade C1, collected between 2005 and 2016 (**II**; Fig. 1A). The phylogeny of the 3D region is seen clearly as more disorganized, with noticeable changes from the clades assigned to the VP1 tree. The large C1 clade from VP1 is seen breaking into smaller C1.1–C1.4 clades within the 3D tree, with the new clades having a more distinct geographical and temporal grouping (**II**; Fig. 1B). Additionally, new sequences were introduced to some of the clades, such as C1.1 adding Finnish and Dutch sequences to the clade. A notable shift in the topology as compared to the VP1 tree can be seen with C6.1 grouping considerably closer to the splintered C1 clade on the 3D side (**II**; Fig. 1B). Clade C2 can be seen losing members in the 3D tree, which end up being major contributors to the overall disorganization of the 3D phylogeny, as the members end up in their own branches without grouping well with any other larger clade. Clade C3, which contained only Dutch and US isolates from the 1970s in VP1, is also seen splintering into two smaller clades, and notably is being disrupted by Finnish isolates from the late 1990s, and a more recent French isolate from 2013 (**II**; Fig. 1B).

5.2.2 Coxsackievirus A9 receptor binding sites

It has been shown that the specific RGD motif located towards the 3´ end of coxsackievirus A9's VP1 capsid protein has a role in cellular entry. More specifically, the motif recognized and binds to the αVβ3 and αVβ6 integrins (Heikkilä et al., 2009; Shakeel et al., 2013). As the motif is located at the tail-end of the VP1 encoding gene, it was able to be extracted from 64 out of the 75 sequences within the data set. All of the sequences contained the RGD motif, which adds to the evidence of the motif being well conserved within these viruses. Additionally, a number of different mutations were mapped to the areas surrounding the motif. Specifically, L/M mutations at position RGD+1, and L/F mutations and position RGD+4.

Previously, a specific T132R/K mutation in the VP1 protein has been suggested to enable heparin binding in CVA9 through heparan sulfate/heparin class of proteoglycans (HSPG) mediating cellular attachment and infection (Merilahti et al., 2016). A BLAST search of this putative heparin binding site was carried out against all non-redundant enteroviruses to analyze if the site could be found in previously published sequence data. The hits included twelve non-redundant enterovirus types (E3, E5, E6, E7, E11, E12, E16, EVB74, EVB85, EVB93, and EVA119). Further literature searches were able to point majority of the species to previously published studies suggesting heparin binding sites for the species (McLeish et al., 2012). However, EVB85 appeared to be a new species with the putative heparin binding site.

5.2.3 Analysis of clinical symptom data

The availability of clinical symptom data gave an opportunity to examine if specific symptoms could be seen forming any clear cluster in the phylogenies. Especially the position of meningitis was examined this way, as it is one of the most severe results of a CVA9 infection, and a disease to which picornaviruses are a major contributor to.

The complete data set contained 21 cases with meningitis (**II**; Table 1). The majority of the cases positioned themselves to clades C1, C2, C3, C5, and C6 in the VP1 phylogeny, with four meningitis causing isolates not being assigned any particular clade. Comparing the distribution of the meningitis causing cases to the 3D phylogeny, no clear shift in the location is seen. That is, the isolates responsible for these symptoms are still grouped in a similar fashion as they were based on VP1. Furthermore, examining the VP1 and 3D phylogenies individually, there is not clear clustering of meningitis cases based on any one factor that can be seen. Similarly, no conclusions could be made with other symptoms, such as diarrhoea.

5.3 Antibody development against human parechoviruses

5.3.1 Functionality and specificity of the developed Mab-PAR-1 against HPeV-A1 virus and HPeV-A1-VP0 protein

The developed IgG1/κ antibody (Mab-PAR-1) that was chosen for further assays was tested against the prototype HPeV-A1, "Harris", strain, as well as HPeV-A1-VP0-GST protein in Western blot, capture ELISA, and IFA. Mab-PAR-1 was seen binding to approximately 32 kDa (lane 1) and 58 kDa (lane 2) proteins in size, which corresponds to the sizes of viral HPeV-A1 VP0 and recombinant VP0-GST proteins (**III**; Fig. 2). Capture ELISA indicated that the antibody binds to native virion structures (**III**; Fig. 2B). Additionally, IFA demonstrated that the antibody binds specifically to HPeV-A1 (**III**; Fig. 2C).

Specificity of the antibody to HPeV-A1 was further analyzed by conducting binding assays against different HPeV-A1 strains, as well as other enteroviruses. Additionally, other parechovirus types were included to determine whether a broader range of HPeVs could be recognized with the antibody. The assays showed that Mab-PAR-1 was able to recognize six different HPeV types 1–6, and that the reactivity across different HPeV-A1 strains was all successful (**III**; Table 1). The cross-reactivity to other viruses was tested against eight enterovirus (EV) and six rhinovirus (RV) types. These viruses gave a signal response to another pan-enterovirus antibody, but did not respond to Mab-PAR-1, meaning Mab-PAR-1 remained very specific to HPeVs (**III**; Table 1, Fig. 4).

5.3.2 Characterization of phage-displayed derived scFv antibodies against HPeV-A1-VP0

The previously described Mab-PAR-1 monoclonal antibody was selected for the development of broad range detector antibodies from a synthetic scFv phage display antibody library. Two libraries, scFvP and scFvM, were developed that differed in their binding site design and, and underwent three rounds of panning. scFvM has previously been shown to be capable of producing binders against both low molecular weight as well as macromolecular antigens, while scFvP is more restricted to targeting marcomolecules (Huovinen et al., 2013). The last round of panning was performed against purified HPeV-A1 "Harris" particles in parallel with HPeV-A1-VP0 to ensure specific binding to native virus. The genes generated by the enriched libraries were cloned into pLK06FT to generate scFv-AP (bacterial alkaline phosphatase) His6-FLAG tag fusion proteins to be used in ELISA screening. After

panning rounds, a total of 84 clones showing elevated signals were sequenced, and 68 were analyzed further. The 68 clones contained 18 unique sequences, and 17 different heavy chain complementarity determining region 3's (CDR-H3). Three clones were identified as being from the scFvP library, and 15 from the scFvM library. Additionally, six clones were obtained through panning with HPeV-A1 virus particles, and seven through panning with the recombinant VP0 protein. The remaining five clones were found in both pannings. Eleven clones were selected for further IMAC purification, all of which were detected in pannings that included a round on the virus particles. Lastly, the IMAC-purified scFv antibodies were tested with TRF against VP0-GST and GST (**V**; Fig. 1), with an additional immunofluorescence assay against HPeV-A1 to confirm. As a result, three binders named scFv-55, scFv-59, and scFv-71, deriving from the scFvM library were selected for further assays.

5.3.3 scFv antibody binding targets and specificity against HPeV types 1–6

Binding of the three scFv antibodies were tested against HPeV-A1 as well as the recombinant HPeV-A1-VP0-GST protein using sandwich ELISA, Western blot, and IFA. HPeV-A1 and VP0 were recognized by all three scFv antibodies, while binding to cell lysate or GST that acted as controls was negative (**V**; Fig. 2a). Western blot showed that all of the scFv antibodies recognized a protein of approximately 37 kDa, which is equivalent to the size of native HPeV-A1-VP0. Additionally, protein with size of 64 kDa was recognized, corresponding to VP0-GST fusion protein (**V**; Fig. 2b). In IFA assays, the scFv antibodies successfully recognized HPeV-A1 "Harris", similarly to the anti-HPeV-A1 polyclonal antiserum acting as a positive control (**V**; Fig. 2c) (Tripathi et al., 2021).

As with the previous Mab-PAR-1 antibody, the scFv antibodies were tested further against HPeV types 1–6, as well as other enteroviruses using IFA. scFv-55 and scFv-71 recognized all of the HPeV types, which was expected as their sequences only differed at five positions, and both possessed a short CDR-H3 loop (**V**; Table 1). On the other hand, scFv-59 did not recognize one of the HPeV-A3 isolates, and also failed to detect HPeV-A4 and HPeV-A5 (**V**; Table 2). Cross-reactivity with other enterovirus types was tested in IFA with eight other enteroviruses together with a pan-enterovirus antibody control. None of the scFv antibodies recognized the other enterovirus types, which was not surprising due to the clear differences present in parechovirus VP0 and other enteroviral proteins.

5.3.4 Identification of the Mab-PAR-1 and scFv antibody structural epitope

In **Publication III**, a 15-mer peptide library was generated using the HPeV-A1-VP0 protein as a model. Subsequent Pepscan's CLIPS analysis identified two signal peaks corresponding to low and high affinity epitopes within VP0. The low affinity epitope was located at 140-ELPKVFWDHQDK-151, while the high affinity epitope was located at 184-LVVYEPKPVVTYDSKLEFGAFT-205. Further refinement of the signal data narrowed the high affinity epitope to 192-VVTYDSKL-199 of VP0, referred to as the "core" (**V**; Fig. 5a). Sequence analysis of 25 HPeV isolates from **Publication III** included types representing HPeV-A1, and HPeV-A3–6. Sequence assembly resulted in 24 usable assemblies, with HPeV-A2 "Williamson" not yielding a sequence. All of the sequenced isolates were previously detected with the described Mab-PAR-1. The low affinity epitope was found specifically in one of the sequences, HPeV-A1 "Harris", and the pairwise amino acid identity of this epitope across the data set was 86.7%. Further analyses were focused on the high affinity epitope, where the exact high affinity epitope core was found in 17 out of 24 sequences, all belonging to HPeV-A1. Overall amino acid identity of the epitope was 76.3% across all sequences, and the remaining sequences contained V192A, T194D, Y195H, D196S/E, S197D/T, and L199M mutations.

A more comprehensive sequence analysis of the high affinity epitope was also carried out pulling all available parechovirus VP0 data from GenBank, including the previously missing HPeV-A2 "Williamson", resulting in a data set of 385 sequences. The sequence identity of the high affinity epitope core across the final data set was surprisingly low at 50%, with the exact sequence found in 107 sequences. Vast majority (n=103) of these sequences belonged to HPeV-A1, while the remaining four belonged to HPeV-A3, HPeV-A8, HPeV-A17, and one untyped HPeV-A. Taking a closer look at the epitope core region identities, it was clear that the epitope was much more conserved within different HPeV types. Focusing on HPeV types 1–6, the intraspecies identities ranged between 82.1% to 100%. It should be noted that the sequence data set is biased towards HPeV-A1 and HPeV-A3 in data availability, as these parechovirus types are the most prevalent and/or pathogenetic in terms of human disease. In these two well represented types, the epitope sequence identities were 94.7% and 97%, respectively. Sequence analysis of the wider epitope area as determined by the Pepscan's CLIPS analysis, confined to the range 184-205 in VP0, showed that the short regions directly flanking the epitope core were very highly conserved even across all HPeV types 1–6 (**V**; Fig. 5b)

Structural analysis of the epitope was performed against HPeV-A1 "Harris" crystal structure (PDB ID: 4Z92). The analysis showed that the epitope core was located mostly on the surface on the capsid with 7 out of the 8 being accessible from the cytoplasmic side of the capsid. Additionally, the epitope was also situated at a

close proximity of the junction between all of the capsid proteins VP0, VP1, and VP3 (**V**; Fig. 5a), and interestingly also shares close proximity to the RGD motif found at the 3´ end of VP1. The fully conserved flanking regions 189-PKP-191 and 200-EFG-202 of the core epitope were also exposed on the capsid surface. The structure shows the epitope core forming a protruding helical loop structure, while the flanking regions appear to stabilize it (**III**; Fig. 5C).

6 Discussion

6.1 The challenge of seemingly simple viruses

Picornaviruses appear as a relatively simple research target at first glance. They are non-enveloped, positive-sense single-stranded RNA viruses with a non-segmented, easy to comprehend genome structure. However, with some of these previously mentioned attributes also come very big challenges when it comes to picornavirus research. Small, positive-sense single-stranded RNA viruses are known to have very high mutation rates due to their naturally error-prone replication process. Combined with the frequent recombination taking place in picornaviruses, it creates a virus population that is constantly evolving to increase its fitness in infecting new hosts and evading immune response, while at the same time being notoriously difficult to develop widely working treatments against. While the issues facing picornavirus research are not unique to picornaviruses themselves, but also concern other RNA viruses, picornaviruses do act as a good model of an RNA virus in this sense. As such, picornaviruses have the potential to aid other RNA virus research as well. From a medical standpoint, while picornaviruses might not be as threatening as some more hostile viruses infecting humans, picornaviruses have still established themselves as one of the most prevalent human pathogens. Frequent epidemics, the high healthcare cost of less severe infections, but still possessing the ability to cause life-threatening diseases, and especially being threatening to small children, picornaviruses have established themselves as pathogens that are waging a war of attrition against humans (De Crom et al., 2016; Monto et al., 2001; Shee & Weber, 2017).

The general aim of this dissertation was to gather a comprehensive view of modern bioinformatics in the setting of everyday picornavirus research. The goal was to widely utilize bioinformatics, and at the same time consider the state of bioinformatics when it comes to virological research. An especially important aspect of the work is to emphasize the use of approachable bioinformatic techniques that can help guide research and analyze results better. Utilization of bioinformatics faces the risk of getting lost in the "flood" of available data. While "Big Data" analysis certainly has its place, the approach might not be suitable for more specific research questions, e.g., in the study of exact infection mechanisms on a molecular level. Picornaviruses suit the task of using bioinformatics in basic virological research well

as a target pathogen, because, as mentioned above, they encompass both simple and very complicated aspects. Additionally, picornaviruses impose a significant burden on healthcare as one of the most prevalent pathogens infecting humans.

Bioinformatics itself is still a relatively young field, although technology and bioinformatic methods are advancing in large leaps. Modern research is often multi-disciplinary and requires the expertise of researchers covering many scientific fields. Bioinformatics itself is no exception in a sense that it combines the computational side with the biological side of science. As such, bioinformaticians should aim to understand both so that new tools can be developed with the best understanding of the biological problem at hand. Likewise, it should be of interest to biologists to take up the use of even basic bioinformatics in their research. While complex data analyses and the building of analysis pipelines may require the expertise of a dedicated bioinformatician, rudimentary sequence and structural analyses could be carried out by anyone with some initial guidance. Moreover, the as biologists also understand more of the bioinformatics side, it enables better communication between scientists to tackle the scientific problem in question.

6.2 Picornaviruses as therapeutic agents or vectors

Picornaviruses are a naturally suitable target to use as a therapeutic agent or as a vector, as they are naturally cytolytic and possess many different infection routes through a wide cell receptor usage (Baggen et al., 2018; Lin et al., 2009; Rossmann et al., 2002). Previously, some enteroviruses have been suggested to fit this role and have in fact shown promise against different cancer types in clinical trials (McCarthy et al., 2019; Ylä-Pelto et al., 2016). One of the common examples of a specific picornavirus used as such is CAVATAK®, which is a cell-adapted coxsackievirus A21. The function of CAVATAK® is based on the increased affinity to the DAF cell surface receptor it has acquired through cell adaptation, which allows it to better infect the specified target cells (Au et al., 2007; Berry et al., 2008; Shafren et al., 2004).

In this dissertation, another oncolytic virotherapy drug called Rigvir® was the subject of two studies. As with CAVATAK®, Rigvir® is also a cell adapted virotherapy virus. Specifically, Rigvir® has been adapted on melanoma cells from a clinical E7 isolate, and belongs to the Enterovirus B species. Likewise, DAF has been previously established to play a role in E7 infection and cellular entry mechanisms, acting as a receptor for clathrin-mediated endocytosis (Kim & Bergelson, 2012). Rigvir® was initially approved and registered in Latvia back in 2004, and has shown prolonged survival of patients suffering from melanoma stage IV M1c, small-cell lung cancer stage IIIA, and histiocytic sarcoma stage IV in a

limited number of case studies (Alberts et al., 2016; Doniņa et al., 2015). However, worryingly, the background information and previous clinical data regarding Rigvir® is dubious, with problems such as proper controls to clinical E7 samples, or the information not being credible by western research standards. Furthermore, Rigvir®'s approval was revoked in 2019 by Latvia's State Agency of Medicines due to discrepancies with laboratory testing of Rigvir® and previously reported results. Regardless, Rigvir®'s marketing is ongoing, and it still holds a license in several countries.

In **Publications I** and **IV**, we subjected Rigvir® to comprehensive analyses against clinical E7 isolates. Bioinformatics analyses were used to determine potential modes of action for Rigvir®'s claimed oncolytic and oncotropic properties. The sequence analysis of Rigvir® as compared to clinical E7 isolates revealed that Rigvir®'s capsid proteins contained some unique mutations. Additionally, one of the mutations, E/Q/N162G, was located directly on the residues previously mapped to form contact with E7 capsid and DAF (Plevka et al., 2010). Overall, the sequence analysis of the DAF contact region of E7 showed a degree of flexibility as judged by the sequence variability (**IV**; Fig. 2). This could potentially have an impact on the virus' ability to adapt to new cell surface receptors, but can also indicate a degree of durability towards mutations while still maintaining adequate binding to DAF. Phylogenetic analysis of Rigvir® revealed that it was the most divergent isolate compared to the other full-length E7 isolates obtained from GenBank. However, this is not particularly surprising considering the nature of Rigvir®'s development, where it has been purposefully passaged on melanoma cells numerous times in an attempt to increase its fitness towards infecting melanoma cells.

Subsequent cellular assays rigorously tested Rigvir® against clinical E7 isolates across two native non-cancer cell lines, and six cancer cell lines. The virus samples were standardized through establishing a TCID50, and assays were performed to observe cytopathic effect to determine the viability of the cells against infection. Additionally, IFA was carried out to enable a clearer view of the infection within cells, and to enable more accurate image analysis. Lastly, the virus amounts were studied with RT-qPCR to determine if replication characteristics differ between Rigvir® and the clinical isolates in a significant manner.

Infection assays described above showed that Rigvir® does not conclusively differ in its infection profile when compared to clinical isolates. Rigvir® did successfully infect and destroy cancer cell lines. However, the same effect was observed in the native isolates, which was expected due to the known naturally cytolytic properties of echoviruses. At the same time, similar to the native isolates, Rigvir® also infected the healthy non-cancer cell lines, which is troublesome as Rigvir® is marketed as being oncotropic and safe to use. In the end, although Rigvir® was found to possess at least some interesting mutations in its capsid proteins, and

was seen phylogenetically distant to clinical isolates, the cellular assays carried out do not support the use of Rigvir® as an oncolytic virotherapy when compared to clinical E7 isolates. Moreover, while the bioinformatic analyses did yield interesting points, it should be noted that the E7-DAF contact interface consist of a total of 59 residues (Plevka et al., 2010). Combined with the fact that the residues in the region already showed sequence variability, the effect of a single mutation should not be overestimated. Lastly, bioinformatic analyses did not detect any other major deviations compared to the clinical isolates that could help explain Rigvir®'s proposed mode of action.

6.3 Coxsackievirus A9 phylogenetics, receptor binding, and epidemiology

CVA9 belongs to the Enterovirus genus, and is one of the most pathogenic enterovirus types, causing diseases such as the common cold, febrile rash, aseptic meningitis, pleurodynia, encephalitis, acute flaccid paralysis (paralytic poliomyelitis), and neonatal sepsis-like disease (Palacios & Oberste, 2005; Zhao et al., 2022). With the VP1 capsid protein being the most heterogenous protein, typing of enteroviruses is commonly based on the distance of VP1 proteins observed between different types (Zell et al., 2017). On the other hand, the non-structural proteins of enteroviruses are subject to less evolutionary pressure, and are considered more conserved across virus types. As the analysis recombination events have become an important concerning picornavirus epidemiology, the analysis of potential recombination can be carried out through phylogenetics by observing incongruence in the topologies of two distant genome regions. At the same time, further epidemiological discoveries can be done based on the phylogenetics themselves, with the addition of clinical data.

CVA9 VP1 capsid protein carries a specific RGD motif that has previously been shown to have a role in cellar entry, and is only found in some picornaviruses (Merilahti et al., 2012). The αVβ6 integrin has also been shown to be a high-affinity receptor, and RGD has also been proposed as an attachment factor for integrins. However, CVA9 infection is not entirely dependent on the RGD motif, as other studies have also established RGD-independent infection pathways (Heikkilä et al., 2016; Hughes et al., 1995; Roivainen et al., 1991). Furthermore, other molecules have been suggested to have a role in CVA9 cellular entry (Baeshen, 2014; K. Triantafilou et al., 2002), such as the HS binding site, although clear role of HS binding remains undiscovered (McLeish et al., 2012; Merilahti et al., 2016).

In **Publication II**, a data set of 75 CVA9 isolates collected between 1959 and 2016 were subjected to sequence, recombination, and phylogenetic analyses. Phylogenetic trees were built from the VP1 and 3D encoding regions using the

BEAST software package (Drummond et al., 2012), which simultaneously build time-inferred phylogenies, giving an estimate of the most recent common ancestor (tMRCA) and the evolutionary rate for the isolates. Additionally, the inclusion of clinical symptom data gave the opportunity to examine potential links to CVA9 pathogenicity. Within VP1, six supported bootstrap clades were identified, and the established clades were compared to the phylogeny generated for the 3D region. Incongruence between the phylogenies suggested possible recombination taking place with isolates belonging to the clades C1 and C6 (**II**; Fig. 1). Clade C1 contained relatively modern isolates from Finland, Asia-Pacific region, and the US, collected between 2005 and 2013. Clade C6 contained isolates that were considerably older, collected between 1974 and 1988 primarily from the Americas region. Additional incongruence was detected between Finnish and Dutch isolates from the C2 and C3 clades. The long time period covered by the collected samples enables the study of geographical movement of CVA9 throughout the years, and the phylogenetics suggest continuous recombination taking place within CVA9 as the virus travels around the world. The addition of clinical data did not yield clear results in terms of being able to detect patterns that tie certain symptoms to specific clusters within the phylogenies. It appears that links to pathogenicity must reside elsewhere in the genomes, and the analysis of other areas of the genome with larger supporting clinical symptom data set would be needed to draw better conclusions.

The RGD motif was found to be completely conserved across the isolates, providing further evidence that the motif remains conserved in nature, although it might not necessarily be required for successful infection. The mutations surrounding the RGD motif were analyzed, as they have been shown to enable binding to different integrins based on the residue at specific locations. For example, the leucine flanking RGD at position +1 has been previously found to stabilize integrin binding to the α5β1 and αVβ3 receptors in another picornavirus, the foot-and-mouth disease virus (FMDV). Additionally, other studies have also shown the leucines at RGD + 1 and +4 to be important for stable binding to integrin αVβ6 (DiCara et al., 2008; Jackson et al., 2000). The mutations observed at these potions could thus potentially have an effect on CVA9 integrin binding, although further studies are required to determine the actual effect (**II**; Fig. 2).

The HS binding motif is suggested to be enabled by a specific T/R132K mutation in VP1, which enables symmetry related cluster of positive charges around the 5-fold axis, thus allowing binding to HS on the cell surface (McLeish et al., 2012). The specific mutation was found at the correct location on another previously unknown enterovirus, EVB85, suggesting that these types could use HS as a receptor for cell entry. However, other studies blocking HS binding at this location have still seen successful infection of CVA9, suggesting that perhaps the role of HS in cellular entry

is either not completely understood, or that there are other HS binding sites within the genome that the virus can still utilize (Merilahti et al., 2016).

6.4 Antibody development against human parechoviruses

Current parechovirus diagnosis is generally carried out by performing RT-qPCR. This is due to the parechoviruses exhibiting a wide and overlapping range of symptoms with other pathogens that precise diagnostic methods such as RT-qPCR are needed (Benschop et al., 2008; Renaud et al., 2011). Rapid diagnostics through RT-qPCR have been developed by utilizing multiplexing, but the method remains expensive and often not a viable option for poorer geographic areas where availability or cost can be an issue. Thus, there is a need for high quality antibodies in diagnostics that can recognize a broad range of parechoviruses. Additionally, phage-display derived scFv antibodies are a form of antibody that can be used for diagnostic and imaging purposes that is produced in a bacterial system. These types of fusion antibodies contain both an antigen binding site as well as a marker, often alkaline phosphatase or a fluorescent protein, which allows for the tracking of the viral infection. As such, these scFv antibodies are a valuable tool in virology for different applications such as diagnostics, therapies, and basic research (Ahmad et al., 2012; Xu et al., 2017; Zhang et al., 2019).

In **Publication III**, a broad range monoclonal antibody (Mab-PAR-1) was developed against a HPeV-A1 VP0 recombinant protein that recognized a wide range of different HPeV-A1 strains, as well as other HPeV types 1–6. At the same time, Mab-PAR-1 showed great specificity to only HPeVs, failing to recognize other entero- or rhinoviruses. In **Publication V**, the antibody developed earlier was subjected to further analysis with the aim to create phage-display derived scFv antibodies. The *E. coli* recombinant HPeV-A1-VP0 target underwent three rounds of phage display scFv antibody library panning to allow for the identification of suitable scFv clones. Further ELISA assays resulted in three confirmed scFv candidates scFv-55, -59, and -71, which were subsequently further purified and tested in ELISA, Western blot, and immunofluorescence microscopy against HPeV types 1–6. ScFv-555 and -71 performed the best, recognizing all of the HPeV types, while maintaining high specificity only to parcehoviruses. ScFv-59 did not recognize all HPeV type 3 isolates, and did not recognize the HPeV types 4 and 5.

Pepscan's CLIPS analysis was used to determine the epitope for the developed antibodies, and narrowed the high affinity epitope to the VP0 region 184-LVVYEPKPVVTYDSKLEFGAFT-205, with the strongest signal strength observed at the epitope "core", at location 192-VVTYDSKL-199. Bioinformatic analyses were performed on parechovirus sequence data covering the VP0 region to determine

the conservation of the epitope across HPeV types. Additional structural analysis of the capsid proteins was carried out to determine where specifically the epitope places on the capsid, and what implications the structure has.

Interestingly, the high affinity epitope core had a relatively low amino acid sequence identity across the whole data set at 50% (*n*=385), and the exact epitope sequence was only found in 107 sequences. However, observing sequence conservation at a virus type level, the epitope site showed much higher conservation between different isolates. Looking at the two most prevalent HPeV types, in HPeV-A1, the epitope core sequence identity was 94.7%, while in HPeV-A3 it was 97%. Similarly high sequence identities were seen with other human HPeV types 2, 4, 5, and 6. Structural analysis of the epitope placed it close to the junction of all the capsid proteins (**V**; Fig. 5a), and interesting also at close structural proximity to the RGD motif located towards the 3′ end of VP1. The epitope core formed a helical loop structure that protruded slightly from the surface. An interesting observation was that even though the epitope core itself is rather variable between HPeV types, the tripeptides flanking the region on both sides remain fully conserved, while also being exposed on the surface together with the epitope core (**V**; Fig. 5b). This suggests that these conserved flanking regions could be structurally important in order to stabilize the epitope core, which itself remains flexible. This flexibility of the epitope could also be the reason why the site allows for recognition of several different HPeV types, regardless of the lower conservation at the core, thus forming a conformational epitope.

Previously developed antibodies against parechoviruses have been limited in their ability to recognize a broader range of parechoviruses, or they have simply not been tested against more HPeV types, often limiting to just HPeV-A1 (Abed et al., 2007; Alho et al., 2003; Chang et al., 2015). The pan-parechovirus antibodies developed in these works could have fruitful diagnostic and research potential, as they recognize a broad range of all HPeV-A1 strains, and have the potential to also recognize HPeV types 2–6.

7 Conclusions

This dissertation aimed to utilize bioinformatics in day-to-day picornavirus research. A general goal was to recognize how bioinformatics can enhance traditional virological research with the application of approachable bioinformatic methods. The dissertation work consisted of five publications, which were grouped into three projects with a common theme. These projects concerned the topics of picornaviral molecular evolution, pircornaviruses in virotherapy, picornavirus pathogenesis, and antibody development.

Based on the research carried out in this dissertation, the following conclusions can be made:

1. Picornaviruses can act as a suitable target for virotherapy applications. However, careful consideration is needed when approaching the subject. The continuing lack of basic knowledge available regarding picornavirus infection mechanisms hamper our understanding of meaningful picornavirus-based virotherapy development. Based on the analyses, Rigvir® cannot be advocated as a safe and effective virotherapy agent.
2. Continuous recombination and at times ambiguous receptor usage during picornavirus infection remain a challenge in picornavirus research in the understanding of their exact infection mechanisms as well as in the development of effective treatments against infection.
3. Antibodies developed against parechoviruses in these studies have the potential to act as pan-parechovirus antibodies against the most important HPeV types infecting humans. Additionally, structural bioinformatics help analyze the conformational epitopes of parechoviruses at a more detailed level.
4. Bioinformatics is a useful and often a necessary tool in even basic virological research, and the importance of utilizing bioinformatics in day-to-day virology continues to grow. Bioinformaticians themselves should take interest in developing tools and methods that are approachable by biologists to promote adoption of bioinformatics by everyone in virology.

Acknowledgements

Research is seldom a solitary task, and as such credit for its completion often belongs to more than those whose present it, just as is the case here. This study was carried out as part of the picornavirus group of the Institute of Biomedicine, University of Turku, Finland during the years 2019–2022.

I would like to express my deepest thanks my supervisor Docent Petri Susi for his support and belief in the project. There have been countless interesting conversations throughout my time in this project that have continued to spark my interest in the fields of virology and bioinformatics, not to mention science in general. At the same time, I have learned a lot about academic life and how to navigate it, and I believe we have made do very well with the resources at hand, which is all thanks to Dr. Susi. Thank you for seeing the importance in advocating the use of bioinformatics for everyone in virology.

Secondly, I would like to thank my follow-up committee Docent Riikka Lund and Professor Urpo Lamminmäki for their support and feedback on the project. I am grateful to Docent Tero Ahola and Docent Jussi Hepojoki for agreeing to act as preliminary examiners of this thesis, and for their insightful comments regarding the work, and thanks to Professor Ilkka Julkunen for agreeing to act as the custos. Additional thanks for the University of Turku itself which has been my home for all of my higher education. Especially thanks to the Turku Doctoral Programme of Molecular Medicine (TuDMM) that accepted me as a doctoral candidate in the programme, as well as the Faculty of Medicine staff who have helped me throughout the years. I would also like to express my deepest thanks to Professor Vesa Hytönen for accepting the invitation to be the opponent for this thesis.

A warm thank you to the people working on the Medisiina D 7th floor who have created an enjoyable working environment, and to those who helped and were willing to teach me in the lab while I was still learning the ropes. An especially important thank you also to all of the co-authors of the publications included in this thesis, to whom I owe a great debt for the work they have done. In addition, my thanks to all of the numerous sources of research grants along the way that helped to fund the work included in this thesis, namely Alfred Kordelin Foundation, Turku

University Foundation, The Finnish Cultural Foundation, Foundation for Research on Viral Diseases, TuDMM, and the University of Turku Faculty of Medicine.

Lastly, I want to thank my family and friends for their unwavering support throughout the years.

November, 2022
Eero Hietanen

References

Abed, Y., Wolf, D., Dagan, R., & Boivin, G. (2007). Development of a Serological Assay Based on a Synthetic Peptide Selected from the VP0 Capsid Protein for Detection of Human Parechoviruses. *Journal of Clinical Microbiology*, *45*(6), 2037–2039. https://doi.org/10.1128/JCM.02432-06

Agol, V. I. (2014). Picornavirus Genetics: an Overview. *Molecular Biology of Picornavirus*, 269–284. https://doi.org/10.1128/9781555817916.CH22

Ahmad, Z. A., Yeap, S. K., Ali, A. M., Ho, W. Y., Alitheen, N. B. M., & Hamid, M. (2012). scFv Antibody: Principles and Clinical Application. *Clinical and Developmental Immunology*, *2012*, 15. https://doi.org/10.1155/2012/980250

Alberts, P., Olmane, E., Brokāne, L., Krastiņa, Z., Romanovska, M., Kupčs, K., Isajevs, S., Proboka, G., Erdmanis, R., Nazarovs, J., & Venskus, D. (2016). Long-term treatment with the oncolytic ECHO-7 virus Rigvir of a melanoma stage IV M1c patient, a small cell lung cancer stage IIIA patient, and a histiocytic sarcoma stage IV patient-three case reports. *APMIS*, *124*(10), 896–904. https://doi.org/10.1111/apm.12576

Alberts, P., Tilgase, A., Rasa, A., Bandere, K., & Venskus, D. (2018). The advent of oncolytic virotherapy in oncology: The Rigvir® story. *European Journal of Pharmacology*, *837*, 117–126. https://doi.org/10.1016/j.ejphar.2018.08.042

Alho, A., Marttila, J., Ilonen, J., & Hyypiä, T. (2003). Diagnostic Potential of Parechovirus Capsid Proteins. *Journal of Clinical Microbiology*, *41*(6), 2294–2299. https://doi.org/10.1128/JCM.41.6.2294-2299.2003

Aoki, Y., Abe, A., Ikeda, T., Abiko, C., Mizuta, K., Yamaguchi, I., & Ahiko, T. (2012). An Outbreak of Exanthematous Disease due to Coxsackievirus A9 in a Nursery in Yamagata, Japan, from February to March 2012. *Japanese Journal of Infectious Diseases*, *65*(4), 367–369. https://doi.org/10.7883/yoken.65.367

Arhab, Y., Bulakhov, A. G., Pestova, T. V., & Hellen, C. U. T. (2020). Dissemination of Internal Ribosomal Entry Sites (IRES) Between Viruses by Horizontal Gene Transfer. *Viruses 2020, Vol. 12, Page 612*, *12*(6), 612. https://doi.org/10.3390/V12060612

Arita, M., Karsch-Mizrachi, I., & Cochrane, G. (2021). The international nucleotide sequence database collaboration. *Nucleic Acids Research*, *49*(D1), D121–D124. https://doi.org/10.1093/nar/gkaa967

Au, G. G., Lincz, L. F., Enno, A., & Shafren, D. R. (2007). Oncolytic Coxsackievirus A21 as a novel therapy for multiple myeloma. *British Journal of Haematology*, *137*(2), 133–141. https://doi.org/10.1111/j.1365-2141.2007.06550.x

Baeshen, N. (2014). *Investigation of the molecular basis of heparan sulphate binding by coxsackievirus A9 identifies multiple mechanisms*.

Baggen, J., Thibaut, H. J., Strating, J. R. P. M. P. M., & Van Kuppeveld, F. J. M. M. (2018). The life cycle of non-polio enteroviruses and how to target it. *Nature Reviews Microbiology*, *16*(6), 368–381. https://doi.org/10.1038/s41579-018-0005-4

Baloch, Z., Ikram, A., Hakim, M. S., & Awan, F. M. (2021). The Impact of Mutations on the Pathogenic and Antigenic Activity of SARS-CoV-2 during the First Wave of the COVID-19 Pandemic: A Comprehensive Immunoinformatics Analysis. *Vaccines 2021, Vol. 9, Page 1410*, *9*(12), 1410. https://doi.org/10.3390/VACCINES9121410

Bayat, A. (2002). Science, medicine, and the future: Bioinformatics. *BMJ : British Medical Journal*, *324*(7344), 1018. https://doi.org/10.1136/BMJ.324.7344.1018

Belov, G. A. (2016). Dynamic lipid landscape of picornavirus replication organelles. *Current Opinion in Virology*, *19*, 1–6. https://doi.org/10.1016/J.COVIRO.2016.05.003

Benschop, K., Thomas, X., Serpenti, C., Molenkamp, R., & Wolthers, K. (2008). High Prevalence of Human Parechovirus (HPeV) Genotypes in the Amsterdam Region and Identification of Specific HPeV Variants by Direct Genotyping of Stool Samples. *Journal of Clinical Microbiology*, *46*(12), 3965–3970. https://doi.org/10.1128/JCM.01379-08

Bentley, K., & Evans, D. J. (2018). Mechanisms and consequences of positive-strand RNA virus recombination. *Journal of General Virology*, *99*(10), 1345–1356. https://doi.org/10.1099/JGV.0.001142/CITE/REFWORKS

Bergelson, J. M., Chan, M., Solomon, K. R., St. John, N. F., Lin, H., & Finberg, R. W. (1994). Decay-accelerating factor (CD55), a glycosylphosphatidylinositol-anchored complement regulatory protein, is a receptor for several echoviruses. *Proceedings of the National Academy of Sciences*, *91*(13), 6245–6248. https://doi.org/10.1073/PNAS.91.13.6245

Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. https://doi.org/10.1093/nar/28.1.235

Berry, L. J., Au, G. G., Barry, R. D., & Shafren, D. R. (2008). Potent Oncolytic activity of human enteroviruses against human prostate cancer. *The Prostate*, *68*(6), 577–587. https://doi.org/10.1002/pros.20741

Boonyakiat, Y., Hughes, P. J., Ghazi, F., & Stanway, G. (2001). Arginine-Glycine-Aspartic Acid Motif Is Critical for Human Parechovirus 1 Entry. *Journal of Virology*, *75*(20), 10000–10004. https://doi.org/10.1128/JVI.75.20.10000-10004.2001

Chang, J.-T., Yang, C.-S., Chen, Y.-S., Chen, B.-C., Chiang, A.-J., Chang, Y.-H., Tsai, W.-L., Lin, Y.-S., Chao, D., & Chang, T.-H. (2015). Genome and Infection Characteristics of Human Parechovirus Type 1: The Interplay between Viral Infection and Type I Interferon Antiviral System. *PLOS ONE*, *10*(2), e0116158. https://doi.org/10.1371/journal.pone.0116158

Chen, Y. H., Du, W., Hagemeijer, M. C., Takvorian, P. M., Pau, C., Cali, A., Brantner, C. A., Stempinski, E. S., Connelly, P. S., Ma, H. C., Jiang, P., Wimmer, E., Altan-Bonnet, G., & Altan-Bonnet, N. (2015). Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses. *Cell*, *160*(4), 619–630. https://doi.org/10.1016/J.CELL.2015.01.032

Cifuente, J. O., & Moratorio, G. (2019). Evolutionary and Structural Overview of Human Picornavirus Capsid Antibody Evasion. *Frontiers in Cellular and Infection Microbiology*, *9*, 283. https://doi.org/10.3389/fcimb.2019.00283

Cui, A., Yu, D., Zhu, Z., Meng, L., Li, H., Liu, J., Liu, G., Mao, N., & Xu, W. (2010). An outbreak of aseptic meningitis caused by coxsackievirus A9 in Gansu, the People's Republic of China. *Virology Journal*, *7*(1), 72. https://doi.org/10.1186/1743-422X-7-72

Cummins, C., Ahamed, A., Aslam, R., Burgin, J., Devraj, R., Edbali, O., Gupta, Di., Harrison, P. W., Haseeb, M., Holt, S., Ibrahim, T., Ivanov, E., Jayathilaka, S., Kadhirvelu, V., Kay, S., Kumar, M., Lathi, A., Leinonen, R., Madeira, F., … Cochrane, G. (2022). The European Nucleotide Archive in 2021. *Nucleic Acids Research*, *50*(D1), D106–D110. https://doi.org/10.1093/NAR/GKAB1051

De Crom, S. C. M. M., Rossen, J. W. A. A., Van Furth, & A. M., Obihara, C. C., van Furth, A. M., & Obihara, C. C. (2016). Enterovirus and parechovirus infection in children: a brief overview. *Eur J Pediatr*, *175*(8), 1023–1029. https://doi.org/10.1007/s00431-016-2725-7

DiCara, D., Burman, A., Clark, S., Berryman, S., Howard, M. J., Hart, I. R., Marshall, J. F., & Jackson, T. (2008). Foot-and-mouth disease virus forms a highly stable, EDTA-resistant complex with its principal receptor, integrin αVβ6: implications for infectiousness. *Journal of Virology*, *82*(3), 1537–1546. https://doi.org/10.1128/JVI.01480-07

Domingo, E., & Perales, C. (2019). Viral quasispecies. *PLoS Genetics*, *15*(10). https://doi.org/10.1371/JOURNAL.PGEN.1008271

Doniņa, S., Strēle, I., Proboka, G., Auziņš, J., Alberts, P. P., Jonsson, B., Venskus, D., Muceniece, A., Doniņa, S., Strele, I., Proboka, G., Auziņš, J., Alberts, P. P., Jonsson, B., Venskus, D., & Muceniece, A. (2015). Adapted ECHO-7 virus Rigvir immunotherapy (oncolytic virotherapy) prolongs survival in melanoma patients after surgical excision of the tumour in a retrospective study. *Melanoma Research*, *25*(5), 421–426. https://doi.org/10.1097/CMR.0000000000000180

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Duffy, S. (2018). Why are RNA virus mutation rates so damn high? *PLOS Biology*, *16*(8), e3000003. https://doi.org/10.1371/JOURNAL.PBIO.3000003

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Etingov, I., Itah, R., Mincberg, M., Keren-Naus, A., Nam, H. J., Agbandje-McKenna, M., & Davis, C. (2008). An extension of the Minute Virus of Mice tissue tropism. *Virology*, *379*(2), 245–255. https://doi.org/10.1016/J.VIROL.2008.06.042

Ferrer-Orta, C., Ferrero, D., & Verdaguer, N. (2015). RNA-Dependent RNA Polymerases of Picornaviruses: From the Structure to Regulatory Mechanisms. *Viruses 2015, Vol. 7, Pages 4438-4460*, *7*(8), 4438–4460. https://doi.org/10.3390/V7082829

Francisco-Velilla, R., Embarc-Buh, A., Abellan, S., & Martinez-Salas, E. (2022). Picornavirus translation strategies. *FEBS Open Bio*, *12*(6), 1125. https://doi.org/10.1002/2211-5463.13400

Furione, M., Guillot, S., Otelea, D., Balanant, J., Candrea, A., & Crainic, R. (1993). Polioviruses with natural recombinant genomes isolated from vaccine-associated paralytic poliomyelitis. *Virology*, *196*(1), 199–208. http://www.ncbi.nlm.nih.gov/pubmed/8102826

Gauthier, J., Vincent, A. T., Charette, S. J., & Derome, N. (2019). A brief history of bioinformatics. *Briefings in Bioinformatics*, *20*(6), 1981–1996. https://doi.org/10.1093/BIB/BBY063

Geneious Prime 2021.0.3. (n.d.). Retrieved October 1, 2022, from http://www.geneious.com/

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, *27*(2), 221–224. https://doi.org/10.1093/molbev/msp259

Gruenert, D. C., Basbaum, C. B., Welsh, M. J., Li, M., Finkbeiner, W. E., & Nadel, J. A. (1988). Characterization of human tracheal epithelial cells transformed by an origin-defective simian virus 40. *Proceedings of the National Academy of Sciences*, *85*(16), 5951–5955. https://doi.org/10.1073/pnas.85.16.5951

Guo, C., McDowell, I. C., Nodzenski, M., Scholtens, D. M., Allen, A. S., Lowe, W. L., & Reddy, T. E. (2017). Transversions have larger regulatory effects than transitions. *BMC Genomics*, *18*(1), 1. https://doi.org/10.1186/S12864-017-3785-4/FIGURES/4

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic Acids Symposium Series* (Vol. 41, pp. 95–98). https://doi.org/citeulike-article-id:691774

Heikkilä, O., Merilahti, P., Hakanen, M., Karelehto, E., Alanko, J., Sukki, M., Kiljunen, S., & Susi, P. (2016). Integrins are not essential for entry of coxsackievirus A9 into SW480 human colon adenocarcinoma cells. *Virology Journal*, *13*(1), 171. https://doi.org/10.1186/s12985-016-0619-y

Heikkilä, O., Susi, P., Stanway, G., & Hyypiä, T. (2009). Integrin αVβ6 is a high-affinity receptor for coxsackievirus A9. *Journal of General Virology*, *90*(1), 197–204. https://doi.org/10.1099/vir.0.004838-0

Hufsky, F., Ibrahim, B., Beer, M., Deng, L., Mercier, P. Le, McMahon, D. P., Palmarini, M., Thiel, V., & Marz, M. (2018). Virologists—Heroes need weapons. *PLOS Pathogens*, *14*(2), e1006771. https://doi.org/10.1371/JOURNAL.PPAT.1006771

Hughes, P. J., Horsnell, C., Hyypiä, T., & Stanway, G. (1995). The coxsackievirus A9 RGD motif is not essential for virus viability. *Journal of Virology*, *69*(12), 8035–8040.

Hughes, P. J., & Stanway, G. (2000). The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation. *Journal of General Virology*, *81*(1), 201–207. https://doi.org/10.1099/0022-1317-81-1-201/CITE/REFWORKS

Huovinen, T., Syrjanpaa, M., Sanmark, H., Brockmann, E.-C., Azhayev, A., Wang, Q., Vehniainen, M., & Lamminmaki, U. (2013). Two ScFv antibody libraries derived from identical VL-VH framework with different binding site designs display distinct binding profiles. *Protein Engineering Design and Selection*, *26*(10), 683–693. https://doi.org/10.1093/protein/gzt037

Ioannou, M., & Stanway, G. (2021). Tropism of Coxsackie virus A9 depends on the +1 position of the RGD (arginine- glycine- aspartic acid) motif found at the C' terminus of its VP1 capsid protein. *Virus Research*, *294*, 198292. https://doi.org/10.1016/j.virusres.2020.198292

Jackson, T., & Belsham, G. J. (2021). Picornaviruses: A View from 3A. *Viruses 2021, Vol. 13, Page 456*, *13*(3), 456. https://doi.org/10.3390/V13030456

Jackson, T., Blakemore, W., Newman, J. W. I., Knowles, N. J., Mould, A. P., Humphries, M. J., & King, A. M. Q. (2000). Foot-and-mouth disease virus is a ligand for the high-affinity binding conformation of integrin αVβ1: Influence of the leucine residue within the RGDL motif on selectivity of integrin binding. *Journal of General Virology*, *81*(5), 1383–1391. https://doi.org/10.1099/0022-1317-81-5-1383

Jiang, W., & Tang, L. (2017). Atomic cryo-EM structures of viruses. *Current Opinion in Structural Biology*, *46*, 122–129. https://doi.org/10.1016/J.SBI.2017.07.002

Johnson, A. D. (2010). An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics*, *26*(10), 1386. https://doi.org/10.1093/BIOINFORMATICS/BTQ098

Joki-Korpela, P., & Hyypiä, T. (1998). Diagnosis and Epidemiology of Echovirus 22 Infections. *Clinical Infectious Diseases*, *27*(1), 129–136. https://doi.org/10.1086/514615

Joki-Korpela, P., Marjomäki, V., Krogerus, C., Heino, J., & Hyypiä, T. (2001). Entry of Human Parechovirus 1. *Journal of Virology*, *75*(4), 1958. https://doi.org/10.1128/JVI.75.4.1958-1967.2001

Jukes, T. H., & Osawa, S. (1993). Evolutionary changes in the genetic code. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *106*(3), 489–494. https://doi.org/10.1016/0305-0491(93)90122-L

Kalynych, S., Pálková, L., & Plevka, P. (2016). The Structure of Human Parechovirus 1 Reveals an Association of the RNA Genome with the Capsid. *Journal of Virology*, *90*(3), 1377–1386. https://doi.org/10.1128/JVI.02346-15

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kim, C., & Bergelson, J. M. (2012). Echovirus 7 Entry into Polarized Intestinal Epithelial Cells Requires Clathrin and Rab7. *MBio*, *3*(2), 1–10. https://doi.org/10.1128/mBio.00304-11

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120. https://doi.org/10.1007/BF01731581

Kolehmainen, P., Oikarinen, S., Koskiniemi, M., Simell, O., Ilonen, J., Knip, M., Hyöty, H., & Tauriainen, S. (2012). Human parechoviruses are frequently detected in stool of healthy Finnish children. *Journal of Clinical Virology*, *54*(2), 156–161. https://doi.org/10.1016/J.JCV.2012.02.006

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/MOLBEV/MSW054

Lai, J. K. F., Sam, I. C., & Chan, Y. F. (2016). The Autophagic Machinery in Enterovirus Infection. *Viruses 2016, Vol. 8, Page 32*, *8*(2), 32. https://doi.org/10.3390/V8020032

Lee, M. S., Chiang, P. S., Luo, S. T., Huang, M. L., Liou, G. Y., Tsao, K. C., & Lin, T. Y. (2012). Incidence rates of enterovirus 71 infections in young children during a nationwide epidemic in

Taiwan, 2008-09. *PLoS Neglected Tropical Diseases*, *6*(2), 6–11. https://doi.org/10.1371/journal.pntd.0001476

Li, Z., Zou, Z., Jiang, Z., Huang, X., & Liu, Q. (2019). Biological Function and Application of Picornaviral 2B Protein: A New Target for Antiviral Drug Development. *Viruses 2019, Vol. 11, Page 510*, *11*(6), 510. https://doi.org/10.3390/V11060510

Lin, J.-Y., Chen, T.-C., Weng, K.-F., Chang, S.-C., Chen, L.-L., & Shih, S.-R. (2009). Viral and host proteins involved in picornavirus life cycle. *Journal of Biomedical Science*, *16*(1), 103. https://doi.org/10.1186/1423-0127-16-103

Lönnrot, M., Sjöroos, M., Salminen, K., Maaronen, M., Hyypiä, T., Hyöty, H., Lönnrot, M., Sjöroos, M., Salminen, K., Maaronen, M., Hyypiä, T., & Hyöty, H. (1999). Diagnosis of enterovirus and rhinovirus infections by RT-PCR and time-resolved fluorometry with lanthanide chelate labeled probes. *Journal of Medical Virology*, *59*(3), 378–384. https://doi.org/10.1002/(SICI)1096-9071(199911)59:3<378::AID-JMV19>3.0.CO;2-I

Lowen, A. C. (2018). It's in the mix: Reassortment of segmented viral genomes. *PLoS Pathogens*, *14*(9). https://doi.org/10.1371/JOURNAL.PPAT.1007200

Lukashev, A. N. (2010). Recombination among picornaviruses. *Reviews in Medical Virology*, *20*(5), 327–337. https://doi.org/10.1002/RMV.660

Mahdieh, N., & Rabbani, B. (2013). An Overview of Mutation Detection Methods in Genetic Disorders. *Iranian Journal of Pediatrics*, *23*(4), 375. /pmc/articles/PMC3883366/

Manning, V. A., Hamilton, S. M., Karplus, P. A., & Ciuffetti, L. M. (2008). The Arg-Gly-Asp-containing, solvent-exposed loop of Ptr ToxA is required for internalization. *Molecular Plant-Microbe Interactions*, *21*(3), 315–325. https://doi.org/10.1094/MPMI-21-3-0315

Martin, G., & Gandon, S. (2010). Lethal mutagenesis and evolutionary epidemiology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1548), 1953. https://doi.org/10.1098/RSTB.2010.0058

Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G., & Diaz-Toledano, R. (2015). Picornavirus IRES elements: RNA structure and host protein interactions. *Virus Research*, *206*, 62–73. https://doi.org/10.1016/j.virusres.2015.01.012

Martino, T. A., Petric, M., Weingartl, H., Bergelson, J. M., Opavsky, M. A., Richardson, C. D., Modlin, J. F., Finberg, R. W., Kain, K. C., Willis, N., Gauntt, C. J., & Liu, P. P. (2000). The Coxsackie-Adenovirus Receptor (CAR) Is Used by Reference Strains and Clinical Isolates Representing All Six Serotypes of Coxsackievirus Group B and by Swine Vesicular Disease Virus. *Virology*, *271*(1), 99–108. https://doi.org/10.1006/VIRO.2000.0324

Marz, M., Beerenwinkel, N., Drosten, C., Fricke, M., Frishman, D., Hofacker, I. L., Hoffmann, D., Middendorf, M., Rattei, T., Stadler, P. F., & Töpfer, A. (2014). Challenges in RNA virus bioinformatics. *Bioinformatics*, *30*(13), 1793–1799. https://doi.org/10.1093/BIOINFORMATICS/BTU105

Mashima, J., Kodama, Y., Fujisawa, T., Katayama, T., Okuda, Y., Kaminuma, E., Ogasawara, O., Okubo, K., Nakamura, Y., & Takagi, T. (2017). DNA Data Bank of Japan. *Nucleic Acids Research*, *45*(D1), D25–D31. https://doi.org/10.1093/NAR/GKW1001

Mason, P. W., Rieder, E., & Baxt, B. (1994). RGD sequence of foot-and-mouth disease virus is essential for infecting cells via the natural receptor but can be bypassed by an antibody-dependent enhancement pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(5), 1932–1936. https://doi.org/10.1073/pnas.91.5.1932

McCarthy, C., Jayawardena, N., Burga, L. N., & Bostina, M. (2019). Developing Picornaviruses for Cancer Therapy. *Cancers*, *11*(5), 685. https://doi.org/10.3390/cancers11050685

McLeish, N. J., Williams, C. H., Kaloudas, D., Roivainen, M. M., & Stanway, G. (2012). Symmetry-related clustering of positive charges is a common mechanism for heparan sulfate binding in enteroviruses. *Journal of Virology*, *86*(20), 11163–11170. https://doi.org/10.1128/JVI.00640-12

McWilliam Leitch, E. C., Harvala, H., Robertson, I., Ubillos, I., Templeton, K., & Simmonds, P. (2009). Direct identification of human enterovirus serotypes in cerebrospinal fluid by amplification and

sequencing of the VP1 region. *Journal of Clinical Virology*, *44*(2), 119–124. https://doi.org/10.1016/j.jcv.2008.11.015

Merilahti, P., Karelehto, E., & Susi, P. (2016). Role of Heparan Sulfate in Cellular Infection of Integrin-Binding Coxsackievirus A9 and Human Parechovirus 1 Isolates. *PLoS ONE*, *11*(1), e0147168. https://doi.org/10.1371/journal.pone.0147168

Merilahti, P., Koskinen, S., Heikkilä, O., Karelehto, E., & Susi, P. (2012). Endocytosis of integrin-binding human picornaviruses. *Advances in Virology*, *2012*. https://doi.org/10.1155/2012/547530

Monto, A. S., Fendrick, A. M., & Sarnes, M. W. (2001). Respiratory illness caused by picornavirus infection: a review of clinical outcomes. *Clinical Therapeutics*, *23*(10), 1615–1627. https://doi.org/10.1016/S0149-2918(01)80133-8

Muehlenbachs, A., Bhatnagar, J., & Zaki, S. R. (2015). Tissue tropism, pathology and pathogenesis of enterovirus infection. *The Journal of Pathology*, *235*(2), 217–228. https://doi.org/10.1002/PATH.4438

Nielsen, A. C. Y., Gyhrs, M. L., Nielsen, L. P., Pedersen, C., & Böttiger, B. (2013). Gastroenteritis and the novel picornaviruses aichi virus, cosavirus, saffold virus, and salivirus in young children. *Journal of Clinical Virology*, *57*(3), 239–242. https://doi.org/10.1016/J.JCV.2013.03.015

Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., German Grehov, O. G., Kandrov, D., Rasputin, K., Syabro, M., & Tleukenov, T. (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Palacios, G., & Oberste, M. (2005). Enteroviruses as agents of emerging infectious diseases. *Journal of Neurovirology*, *11*(5), 424–433. https://doi.org/10.1080/13550280591002531

Paul, A. V., & Wimmer, E. (2015). Initiation of protein-primed picornavirus RNA synthesis. *Virus Research*, *206*, 12–26. https://doi.org/10.1016/J.VIRUSRES.2014.12.028

Peersen, O. B. (2017). Picornaviral polymerase structure, function, and fidelity modulation. *Virus Research*, *234*, 4–20. https://doi.org/10.1016/J.VIRUSRES.2017.01.026

Peltola, V., Waris, M., Osterback, R., Susi, P., Ruuskanen, O., Hyypiä, T., Österback, R., Susi, P., Ruuskanen, O., & Hyypiä, T. (2008). Rhinovirus Transmission within Families with Children: Incidence of Symptomatic and Asymptomatic Infections. *The Journal of Infectious Diseases*, *197*(3), 382–389. https://doi.org/10.1086/525542

Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., & González-Candelas, F. (2015). Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. *Infection, Genetics and Evolution*, *30*, 296. https://doi.org/10.1016/J.MEEGID.2014.12.022

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Phadke, S., Macherla, S., & Scheuermann, R. H. (2021). Database and Analytical Resources for Viral Research Community. In *Encyclopedia of Virology* (pp. 141–152). Elsevier. https://doi.org/10.1016/B978-0-12-809633-8.20995-3

Plevka, P., Hafenstein, S., Harris, K. G., Cifuente, J. O., Zhang, Y., Bowman, V. D., Chipman, P. R., Bator, C. M., Lin, F., Medof, M. E., & Rossmann, M. G. (2010). Interaction of Decay-Accelerating Factor with Echovirus 7. *Journal of Virology*, *84*(24), 12665–12674. https://doi.org/10.1128/JVI.00837-10

Pons-Salort, M., Parker, E. P. K., & Grassly, N. C. (2015). The epidemiology of non-polio enteroviruses: Recent advances and outstanding questions. In *Current Opinion in Infectious Diseases* (Vol. 28, Issue 5, pp. 479–487). https://doi.org/10.1097/QCO.0000000000000187

Porter, A. G. (1993). Picornavirus nonstructural proteins: emerging roles in virus replication and inhibition of host cell functions. *Journal of Virology*, *67*(12), 6917–6921. https://doi.org/10.1128/JVI.67.12.6917-6921.1993

Puenpa, J., Vongpunsawad, S., Österback, R., Waris, M., Eriksson, E., Albert, J., Midgley, S., Fischer, T. K., Eis-Hübinger, A. M., Cabrerizo, M., Gaunt, E., Simmonds, P., & Poovorawan, Y. (2016).

Molecular epidemiology and the evolution of human coxsackievirus A6. *Journal of General Virology*, *97*(12), 3225–3231. https://doi.org/10.1099/jgv.0.000619

Ren, J., Wang, X., Hu, Z., Gao, Q., Sun, Y., Li, X., Porta, C., Walter, T. S., Gilbert, R. J., Zhao, Y., Axford, D., Williams, M., McAuley, K., Rowlands, D. J., Yin, W., Wang, J., Stuart, D. I., Rao, Z., & Fry, E. E. (2013). Picornavirus uncoating intermediate captured in atomic detail. *Nature Communications 2013 4:1*, *4*(1), 1–7. https://doi.org/10.1038/ncomms2889

Renaud, C., Kuypers, J., Ficken, E., Cent, A., Corey, L., & Englund, J. A. (2011). Introduction of a novel parechovirus RT-PCR clinical test in a regional medical center. *Journal of Clinical Virology*, *51*(1), 50–53. https://doi.org/10.1016/j.jcv.2011.02.010

Roivainen, M., Hyypiä, T., Piirainen, L., Kalkkinen, N., Stanway, G., & Hovi, T. (1991). RGD-dependent entry of coxsackievirus A9 into host cells and its bypass after cleavage of VP1 protein by intestinal proteases. *Journal of Virology*, *65*(9), 4735–4740.

Rossmann, M. G., He, Y., & Kuhn, R. J. (2002). Picornavirus–receptor interactions. *Trends in Microbiology*, *10*(7), 324–331. https://doi.org/10.1016/S0966-842X(02)02383-1

Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cellular and Molecular Life Sciences 2016 73:23*, *73*(23), 4433–4448. https://doi.org/10.1007/S00018-016-2299-6

Santti, J., Harvala, H., Kinnunen, L., & Hyypiä, T. (2000). Molecular epidemiology and evolution of coxsackievirus A9. *Journal of General Virology*, *81*(5), 1361–1372. https://doi.org/10.1099/0022-1317-81-5-1361

Santti, J., Hyypiä, T., Kinnunen, L., & Salminen, M. (1999). Evidence of recombination among enteroviruses. *Journal of Virology*, *73*(10), 8741–8749. http://www.ncbi.nlm.nih.gov/pubmed/10482628

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., & Karsch-Mizrachi, I. (2021). GenBank. *Nucleic Acids Research*, *49*(D1), D92–D96. https://doi.org/10.1093/NAR/GKAA1023

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Shafren, D. R., Au, G. G., Nguyen, T., Newcombe, N. G., Haley, E. S., Beagley, L., Johansson, E. S., Hersey, P., & Barry, R. D. (2004). Systemic Therapy of Malignant Human Melanoma Tumors by a Common Cold-Producing Enterovirus, Coxsackievirus A21. *Clinical Cancer Research*, *10*(1), 53–60. https://doi.org/10.1158/1078-0432.CCR-0690-3

Shafren, D. R., Dorahy, D. J., Greive, S. J., Burns, G. F., & Barry, R. D. (1997). Mouse cells expressing human intercellular adhesion molecule-1 are susceptible to infection by coxsackievirus A21. *Journal of Virology*, *71*(1), 785–789. https://doi.org/10.1128/JVI.71.1.785-789.1997

Shakeel, S., Seitsonen, J. J. T., Kajander, T., Laurinmäki, P., Hyypiä, T., Susi, P., & Butcher, S. J. (2013). Structural and functional analysis of coxsackievirus A9 integrin αvβ6 binding and uncoating. *Journal of Virology*, *87*(7), 3943–3951. https://doi.org/10.1128/JVI.02989-12

Shapiro, B., Rambaut, A., Pybus, O. G., & Holmes, E. C. (2006). A phylogenetic method for detecting positive epistasis in gene sequences and its application to RNA virus evolution. *Molecular Biology and Evolution*, *23*(9), 1724–1730. https://doi.org/10.1093/molbev/msl037

Shastry, B. S., & Shastry, B. S. (2002). SNP alleles in human disease and evolution. *Journal of Human Genetics 2002 47:11*, *47*(11), 561–566. https://doi.org/10.1007/s100380200086

Shee, A., & Weber, H. (2017). Human Parechovirus Infection in Neonates and Children: An Overview. *Journal of Pediatric Infectious Diseases*, *12*(2), 99–103. https://doi.org/10.1055/S-0037-1602384/ID/JR1600034-33

Simon-Loriere, E., & Holmes, E. C. (2011). Why do RNA viruses recombine? *Nature Reviews Microbiology 2011 9:8*, *9*(8), 617–626. https://doi.org/10.1038/nrmicro2614

Slootstra, J. W., Puijk, W. C., Ligtvoet, G. J., Langeveld, J. P. M., & Meloen, R. H. (1996). Structural aspects of antibody-antigen interaction revealed through small random peptide libraries. *Molecular Diversity*, *1*(2), 87–96. https://doi.org/10.1007/BF01721323

Smith, J. B. (2001). Peptide Sequencing by Edman Degradation. In *eLS*. Wiley. https://doi.org/10.1038/npg.els.0002688

Smura, T., Kakkola, L., Blomqvist, S., Klemola, P., Parsons, A., Kallio-Kokko, H., Savolainen-Kopra, C., Kainov, D. E., & Roivainen, M. (2013). Molecular evolution and epidemiology of echovirus 6 in Finland. *Infection, Genetics and Evolution*, *16*(March 2018), 234–247. https://doi.org/10.1016/j.meegid.2013.02.011

Smyth, M. S., & Martin, J. H. (2002). Picornavirus uncoating. *Molecular Pathology*, *55*(4), 214. https://doi.org/10.1136/MP.55.4.214

Smyth, M. S., Pettitt, T., Symonds, A., & Martin, J. (2003). Identification of the pocket factors in a picornavirus. *Archives of Virology 2002 148:6*, *148*(6), 1225–1233. https://doi.org/10.1007/S00705-002-0974-4

Sridhar, Karelehto, Brouwer, Pajkrt, & Wolthers. (2019). Parechovirus A Pathogenesis and the Enigma of Genotype A-3. *Viruses*, *11*(11), 1062. https://doi.org/10.3390/v11111062

Stanway, G., Hovi, T., Knowles, N. J., & Hyypiä, T. (2014). Molecular and Biological Basis of Picornavirus Taxonomy. In *Molecular Biology of Picornavirus* (pp. 15–24). ASM Press. https://doi.org/10.1128/9781555817916.ch2

Steil, B. P., & Barton, D. J. (2009). Cis-active RNA elements (CREs) and picornavirus RNA replication. *Virus Research*, *139*(2), 240–252. https://doi.org/10.1016/J.VIRUSRES.2008.07.027

Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, *18*(1), 100–106. https://doi.org/10.1038/s41592-020-01018-x

Sugauchi, F., Orito, E., Ichida, T., Kato, H., Sakugawa, H., Kakumu, S., Ishida, T., Chutaputti, A., Lai, C. L., Gish, R. G., Ueda, R., Miyakawa, Y., & Mizokami, M. (2003). Epidemiologic and virologic characteristics of hepatitis B virus genotype B having the recombination with genotype C. *Gastroenterology*, *124*(4), 925–932. https://doi.org/10.1053/GAST.2003.50140

Tapparel, C., Siegrist, F., Petty, T. J., & Kaiser, L. (2013). Picornavirus and enterovirus diversity with associated human diseases. *Infection, Genetics and Evolution*, *14*(1), 282–293. https://doi.org/10.1016/J.MEEGID.2012.10.016

Tauriainen, S., Martiskainen, M., Oikarinen, S., Lönnrot, M., Viskari, H., Ilonen, J., Simell, O., Knip, M., & Hyöty, H. (2007). Human parechovirus 1 infections in young children—no association with type 1 diabetes. *Journal of Medical Virology*, *79*(4), 457–462. https://doi.org/10.1002/JMV.20831

Tilgase, A., Patetko, L., Blāķe, I., Ramata-Stunda, A., Borodušķis, M. M., Alberts, P. P., Blaķe, I., Ramata-Stunda, A., Borodušķis, M. M., & Alberts, P. P. (2018). Effect of the oncolytic ECHO-7 virus Rigvir® on the viability of cell lines of human origin in vitro. *Journal of Cancer*, *9*(6), 1033–1049. https://doi.org/10.7150/jca.23242

Triantafilou, K., Fradelizi, D., Wilson, K., & Triantafilou, M. (2002). GRP78, a Coreceptor for Coxsackievirus A9, Interacts with Major Histocompatibility Complex Class I Molecules Which Mediate Virus Internalization. *Journal of Virology*, *76*(2), 633–643. https://doi.org/10.1128/JVI.76.2.633-643.2002

Triantafilou, M., Wilson, K. M., & Triantafilou, K. (2001). Identification of echovirus 1 and coxsackievirus A9 receptor molecules via a novel flow cytometric quantification method. *Cytometry*, *43*(4), 279–289. https://doi.org/10.1002/1097-0320(20010401)43:4<279::AID-CYTO1060>3.0.CO;2-B

Tripathi, L., Hietanen, E., Merilahti, P., Teixido, L., Sanchez-Alberola, N., Tauriainen, S., & Susi, P. (2021). Monoclonal antibody against VP0 recognizes a broad range of human parechoviruses. *Journal of Virological Methods*, *293*, 114167. https://doi.org/10.1016/j.jviromet.2021.114167

Tuthill, T. J., Groppelli, E., Hogle, J. M., & Rowlands, D. J. (2010). *Picornaviruses* (pp. 43–89). Springer, Berlin, Heidelberg. https://doi.org/10.1007/82_2010_37

Wang, Z., & Moult, J. (2001). SNPs, protein structure, and disease. *Human Mutation*, *17*(4), 263–270. https://doi.org/10.1002/HUMU.22

Williams, C. H., Kajander, T., Hyypia, T., Jackson, T., Sheppard, D., & Stanway, G. (2004). Integrin αVβ6 Is an RGD-Dependent Receptor for Coxsackievirus A9. *Journal of Virology*, *78*(13), 6967–6973. https://doi.org/10.1128/JVI.78.13.6967-6973.2004

Winston, D. S., & Boehr, D. D. (2021). The Picornavirus Precursor 3CD Has Different Conformational Dynamics Compared to 3Cpro and 3Dpol in Functionally Relevant Regions. *Viruses 2021, Vol. 13, Page 442*, *13*(3), 442. https://doi.org/10.3390/V13030442

Wolthers, K. C., Susi, P., Jochmans, D., Koskinen, J., Landt, O., Sanchez, N., Palm, K., Neyts, J., & Butcher, S. J. (2019). Progress in human picornavirus research: New findings from the AIROPico consortium. *Antiviral Research*, *161*, 100–107. https://doi.org/10.1016/J.ANTIVIRAL.2018.11.010

Xiang, W., Harris, K. S., Alexander, L., & Wimmer, E. (1995). Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. *Journal of Virology*, *69*(6), 3658–3667. https://doi.org/10.1128/jvi.69.6.3658-3667.1995

Xu, X., Zhang, R., & Chen, X. (2017). Application of a single-chain fragment variable (scFv) antibody for the confirmatory diagnosis of hydatid disease in non-endemic areas. *Electronic Journal of Biotechnology*, *29*, 57–62. https://doi.org/10.1016/J.EJBT.2017.07.003

Yates, J. R., Eng, J. K., McCormack, A. L., & Schieltz, D. (1995). Method to Correlate Tandem Mass Spectra of Modified Peptides to Amino Acid Sequences in the Protein Database. *Analytical Chemistry*, *67*(8), 1426–1436. https://doi.org/10.1021/ac00104a020

Ylä-Pelto, J., Tripathi, L., & Susi, P. (2016). Therapeutic Use of Native and Recombinant Enteroviruses. *Viruses*, *8*(3), 57. https://doi.org/10.3390/v8030057

Yuan, F., Wang, L., Fang, Y., & Wang, L. (2021). Global SNP analysis of 11,183 SARS-CoV-2 strains reveals high genetic diversity. *Transboundary and Emerging Diseases*, *68*(6), 3288–3304. https://doi.org/10.1111/TBED.13931

Zell, R. (2018). Picornaviridae—the ever-growing virus family. *Archives of Virology*, *163*(2), 299–317. https://doi.org/10.1007/s00705-017-3614-8

Zell, R., Delwart, E., Gorbalenya, A. E., Hovi, T., King, A. M. Q., Knowles, N. J., Lindberg, A. M., Pallansch, M. A., Palmenberg, A. C., Reuter, G., Simmonds, P., Skern, T., Stanway, G., & Yamashita, T. (2017). ICTV Virus Taxonomy Profile: Picornaviridae. *Journal of General Virology*, *98*(10), 2421–2422. https://doi.org/10.1099/jgv.0.000911

Zhang, F., Chen, Y., Ke, Y., Zhang, L., Zhang, B., Yang, L., & Zhu, J. (2019). Single Chain Fragment Variable (scFv) Antibodies Targeting the Spike Protein of Porcine Epidemic Diarrhea Virus Provide Protection against Viral Infection in Piglets. *Viruses 2019, Vol. 11, Page 58*, *11*(1), 58. https://doi.org/10.3390/V11010058

Zhao, H., Wang, J., Chen, J., Huang, R., Zhang, Y., Xiao, J., Song, Y., Ji, T., Yang, Q., Zhu, S., Wang, D., Lu, H., Han, Z., Zhang, G., Li, J., & Yan, D. (2022). Molecular Epidemiology and Evolution of Coxsackievirus A9. *Viruses 2022, Vol. 14, Page 822*, *14*(4), 822. https://doi.org/10.3390/V14040822

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2022

ISBN 978-951-29-9081-8 (PRINT)
ISBN 978-951-29-9082-5 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)