A Molecular Approach to Precision Medicine in South African Children with Epilepsy: Towards a Genetics-Based Diagnostic Service for Epilepsy in Childhood

Caitlin McIntosh

MCNCAI001

This dissertation is submitted in fulfilment of the requirements of the degree of a Masters in Medical Science by Research, specialising in Human Genetics.

Division of Human Genetics

Department of Pathology

Faculty of Health Sciences

The University of Cape Town

June 2023

Plagiarism Declaration

1. I, Caitlin McIntosh, know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the American Journal of Human Genetics convention for citation and referencing. Each contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed, and has been cited and referenced. Any section taken from an internet source has been referenced to that source.
3. This dissertation is my own work, and is in my own words (except where I have attributed it to others).
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.
5. I acknowledge that copying someone else's work, or part of it, is wrong, and declare that this is my own work.

Signature: _____ Signed by candidate _____

Acknowledgements

Thank you to the National Research Foundation and the South African Medical Research Council for student and project funding.

Thank you to all study participants and their families for their participation in this study. Thank you to all staff and student volunteers who aided on the parent project in participant recruitment at epilepsy clinics.

Thank you to Dr Gloudi Agenbag, Lerato Majara, Akshay Vanmali, Shrinav Dawlat, Nicole Midgely, and Jessica Chambers for help in the laboratory. Thank you to Professor Gemma Carvill and Professor Jo Wilmshurst for advice and guidance regarding the design and implementation of the epilepsy gene panel. Thank you to Professor Collet Dandara, Dr Marelize Immelman, and Dr Khuthala Mnika for advice on the pharmacogenetics portion of the project.

Finally, thank you to my supervisors, Alina Esterhuizen and Karen Fieggen, for their support, guidance, and patience over the course of this project.

Table of Contents

Abstract

Background:

Epilepsy is a neurological disorder characterised by unprovoked, recurring seizures. Sub-Saharan Africa carries the highest burden of epilepsy in the world, owing mainly to the increased risk factors such as infectious and parasitic disease, and traumatic brain injury. A proportion of this burden is genetic, however, genetic testing for epilepsy in South Africa is limited, currently only accessible in the private healthcare sector via international referral. Genetic testing in epilepsy is an internationally recognised diagnostic tool which can inform the diagnosis and patient management, with options for precision treatment, frequently resulting in improved outcomes. The main aim of this research project, therefore, was to design and validate a next-generation sequencing (NGS) gene panel for South African paediatric patients with drug-resistant epilepsy. The panel design was aimed primarily at the developmental and epileptic encephalopathies, where the yield of informative findings is highest, and the results often carry implications for treatment. The secondary aim of the project was to perform preliminary pharmacogenetic testing, to determine if variants previously associated with antiseizure drug metabolism in other populations are present in this study group.

Materials and Methods:

Forty probands with clinically complex, drug-resistant epilepsy with no identified, acquired cause were recruited from the neurology epilepsy service at the Red Cross War Memorial Children's Hospital in Cape Town. All 40 probands were tested with a panel of 78 genes selected on the basis of association with disease and clinical actionability. NGS data files were subject to variant prioritisation, followed by variant confirmation by cycle sequencing, segregation analysis, and final classification according to ACMG criteria. All 40 probands were tested with two pharmacogenomic arrays, one generalised array, the Veridose® Core Panel produced by Agena Bioscience (San Diego, USA), and one custom-designed anti-seizure medication-specific SNV array targeting eight SNVs across six different genes.

Results:

Three pathogenic variants were identified; two in *SCN1A* and one in *GRIN2A*, with a pickup rate of 7.5% (3/40). Two variants of uncertain significance were identified in *GABRG2* and *GRIN2B*. The findings were aligned with the electroclinical features in each patient. The pharmacogenomic analysis revealed one variant, *EPHX1* rs1051740, which appeared to be statistically significantly differently distributed in the study group to the general African

population (P = 0.02). Three other variants, *EPHX1* rs2234922, *SCN1A* rs3812718, and *CYP2D6* s59421388 appear to be trending towards significance, with P-values of 0.08, 0.05, and 0.05, respectively.

Conclusion:

The main outcome of this project was the successful development of an NGS-based diagnostic protocol for paediatric epilepsy using the gene panel approach, now ready for implementation in the local diagnostic testing laboratory. Moreover, the genetic cause of epilepsy was identified in three study participants, with treatment implications in two (variants in *SCN1A*). The pharmacogenomic analysis in this project did not reveal specific insights owing to the small study group, but provided exposure to pharmacogenomic analysis, and may be used as a basis for further research into the pharmacogenetics of epilepsy in larger African cohorts.

Plan of Dissertation

This dissertation is divided into four chapters. The first chapter provides an introductory background to paediatric epilepsy in Africa, and the means by which a genetic diagnosis for epilepsy may be achieved, as well as brief overview of the pharmacogenetics of epilepsy. It also includes the aims and objectives of the study and the scientific approach to the investigation. The second chapter details the methods and materials of the project's laboratory and bioinformatic processes. The third chapter follows with a presentation of the results of the study. The fourth chapter discusses the findings of the project, the challenges and limitations of the project, and what conclusions were ultimately drawn.

List of Abbreviations

ACE	Active Convulsive Epilepsy
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal Dominant
ADME	Absorption, Distribution, Metabolism, and Excretion
ADMET	Absorption, Distribution, Metabolism, Excretion, and Transport
ADHD	Attention Deficit Hyperactivity Disorder
ADR	Adverse Drug Reaction
AR	Autosomal Recessive
ASD	Autism Spectrum Disorder
ASM	Anti-Seizure Medication
ASO	Anti-Sense Oligonucleotide
BFIS	Benign Familial Infantile Seizures
CMA	Chromosomal Microarray
CNS	Central Nervous System
CNV	Copy Number Variant
CYP	Cytochrome P450
DE	Developmental Encephalopathy
DEE	Developmental Epileptic Encephalopathy
DNA	Deoxyribonucleic Acid
DOC	Depth Of Coverage
DS	Dravet Syndrome
EIEE	Early Infantile Epileptic Encephalopathy
EE	Epileptic Encephalopathy
EIMFS	Epilepsy of Infancy with Migrating Focal Seizures
ES	Exome Sequencing
GOF	Gain-of-Function
GS	Genome Sequencing
HGMD	Human Gene Mutation Database
HIC	High Income Country
HIV	Human Immunodeficiency Virus
HPO	Human Phenotype Ontology

HWE	Hardy-Weinberg Equilibrium
IBE	International Bureau of Epilepsy
IGV	Integrative Genomics Viewer
ILAE	International League Against Epilepsy
INDEL	Insertion/Deletion
ISP	Ion Sphere Particle
LMIC	Low to Middle Income Country
LOF	Loss-of-Function
LOVD	Leiden Open Variation Database
MAF	Minor Allele Frequency
MRI	Magnetic Resonance Imaging
NGS	Next-generation Sequencing
OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase Chain Reaction
PGM	Personal Genome Machine
pM	Picomole
RCWMCH	Red Cross War Memorial Children's Hospital
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SSA	Sub-Saharan Africa
TBI	Traumatic Brain Injury
TSC	Tuberous Sclerosis Complex
UCT	University of Cape Town
VUS	Variant of Uncertain Significance
WHO	World Health Organisation

Chapter 1: Introduction

1.1. Epilepsy in Africa

Epilepsy is estimated to affect up to 70 million people globally, more than 85% of whom live in low to middle income countries (LMICs)[1]. Sub-Saharan Africa (SSA) has a particularly high prevalence of epilepsy, estimated by door-to-door studies to be double that of Europe, North America, and Asia[1]. The increased prevalence of epilepsy in LMICs is largely attributed to increased risk factors, such as the high incidence of central nervous system (CNS) infections, parasitic diseases, traumatic brain injury (TBI), perinatal insult, and limited healthcare services[2].

TBI is most commonly caused by motor vehicle accidents, assault, and injury in war, all of which are more prevalent in SSA than in the rest of the world[3]. Diseases of parasitic origin that may cause or contribute to the development of epilepsy include neurocysticercosis (a common neurological complication of *Taenia solium* infection - a common cause of epilepsy across Africa, Asia, and Latin America), toxocariasis, onchocerciasis (also known as river blindness), and cerebral malaria[3]. The movement of people across different regions and countries means that that healthcare workers may encounter cases of epilepsy caused by diseases not endemic in their own country – such as cases of epilepsy caused by cerebral malaria observed in South Africa.

The distribution of causes of epilepsy in SSA differs from that in high income countries (HICs), as well as between adults and children in SSA (Figure 1). The age groups in which epilepsy is most prevalent also differ between countries, though a particularly high prevalence in individuals younger than twenty years old has been observed in LMICs[3]. Epilepsy of a genetic aetiology is rarely identified in SSA and little is known about its true prevalence or the genetic underpinning of this common condition in the region[4]. This is partly due to misdiagnosis through lack of awareness and assumption of an infectious or other acquired cause, and partly due to limited access to genetic testing[5].

The genetic origin of many epilepsies is often obscured by *de novo* occurrence in the absence of family history (for example, in cases of developmental and epileptic encephalopathy (DEE)).

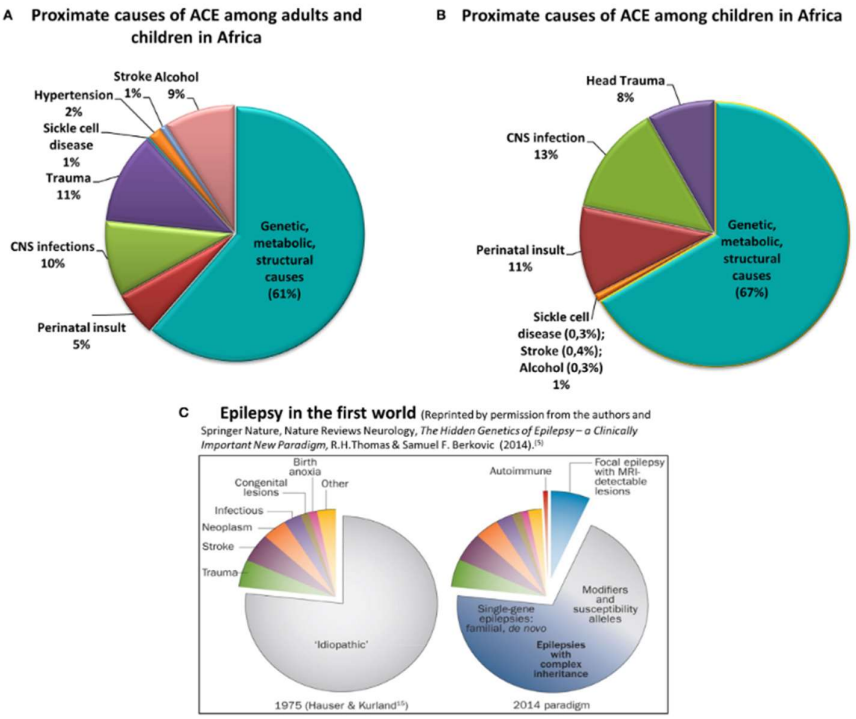


Figure 1 Causes of active convulsive epilepsy (ACE) in **(A)** adults and children in Africa, **(B)** children in Africa, **(C)** all epilepsy cases in the first world. Note the distribution of causes differs between the affected general African population and affected African children, and between the affected general African population and affected individuals in HICs. (Figure adapted from Esterhuizen, et al. 2018.)[4]

However, a family history of epilepsy does not necessarily equate with a genetic aetiology, as it is possible that shared environmental risk factors may cause or contribute to the recurrence of epilepsy among multiple individuals in the same family[3].

1.1.1. Prevalence

The median prevalence of epilepsy in SSA over the past decade has been estimated at 14.2 cases per 1000 people (IQR 8.0 – 33.2), in contrast to 5.8 cases per 1000 people (IQR 2.7 – 12.4) in HICs[4]. The observed prevalence of epilepsy also differs between countries within SSA (Figure 2)[6]. These differences may be a result of incomplete reporting, or different methodological approaches in calculating prevalence, or could be due to differing biological factors (including genetic factors) across the continent[3]. The cause of epilepsy in over 60% of affected children in Africa is unknown, suggesting the possibility of underlying genetic,

metabolic, and structural causes[5]. An accurate estimation of the true prevalence of genetic epilepsies in Africa requires extensive future research.

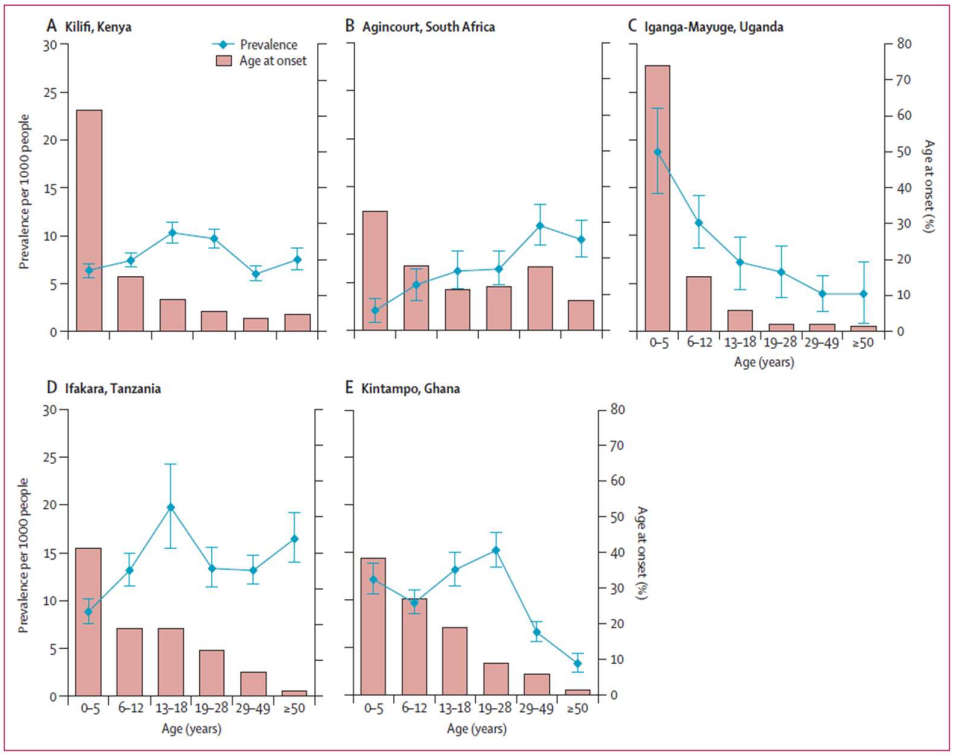


Figure 2 Prevalence of epilepsy across different sites in SSA, in **(A)** Kenya, **(B)** South Africa, **(C)** Uganda, **(D)** Tanzania, and **(E)** Ghana. The blue line graph measuring "Prevalence" indicates an adjusted prevalence to enable comparison between two different survey methods. (Figure adapted from Ngugi et al. 2013.)[6]

1.1.2. Management challenges and consequences

Epilepsy, especially when poorly controlled, is a major public health burden, particularly in countries which already face substantial challenges to their healthcare systems. Effective treatment of epilepsy to mitigate its effects and its co-morbidities is vital. In 2010, a Global Burden of Disease Study found that only HIV infection has a greater disability burden than severe, uncontrolled epilepsy[7].

The impact of the high prevalence of epilepsy in the LMICs is exacerbated by the significant treatment gap in those countries. The treatment gap is defined as the proportion of people who require medical treatment for epilepsy but do not receive it[8]. This gap in the use of anti-seizure medications (ASMs) exists in all contexts, but is especially wide in SSA[3]. This is

primarily due to the shortage of skilled personnel and medical facilities, non-adherence to treatment (sometimes due to indigenous beliefs about the causes of epilepsy), inaccessibility of drugs (due to high costs or unavailability), and the psychosocial effects of the disease, including stigma against sufferers, who are often reluctant to seek medical treatment[3]. Uncontrolled epilepsy can cause severe physical and psychiatric pathologies, with serious consequences such as premature death, physical injuries, such as burns or bone fractures, and psychiatric co-morbidities[3]. Psychiatric co-morbidities associated with epilepsy include increased rates of depression, anxiety, and attention deficit hyperactivity disorder (ADHD), which are observed in children with epilepsy at a rate five times higher than the age-matched general population[9]. The treatment gap is especially relevant in paediatric epilepsy, where correct, early diagnosis and treatment are important in minimising the consequences of seizure activity in the developing brain[4].

Aside from pharmacological treatment of epilepsy with ASMs, surgical interventions for epilepsy are possible. This involves removal of the minimum amount of brain tissue needed to stop the occurrence of seizures (termed the epileptogenic zone) without major complications[10]. Such surgical treatment is becoming increasingly common for intractable epilepsy in regions where it is available[11]. Surgery as treatment for epilepsy in Africa is only available in a few tertiary care centres and private clinics, due to a lack of trained neurosurgeons and necessary infrastructure, as well as limited financial resources[3].

1.1.3. Psychosocial aspects

Aside from co-morbidities of a physical and psychiatric nature, people affected with epilepsy also face social stigma and reduced quality of life[5]. Individuals affected with epilepsy are less likely to be educated or married and are more likely to be unemployed and experience premature mortality than unaffected members of the same population[2,3].

Social stigma and its consequences have been exacerbated in many past and present cultures by epilepsy's strong association with supernatural and paranormal phenomena. In Ancient Greece, epilepsy was often referred to as the "sacred disease" due to its association with numerous deities and mythological figures, hence the title of the medical treatise *On the Sacred Disease*, written by Hippocrates (or an associate) around the end of the 5th century.

Hippocrates called epilepsy “the great disease”, hence the term *grand mal*, a term used to this day to describe a subset of seizures[12]. Epilepsy’s strong supernatural associations have in large part receded in many parts of the world but remain particularly prevalent in Africa, where traditional beliefs remain at the forefront of how epilepsy is perceived, and where epilepsy may be associated with witchcraft and evil spirits, or be believed to be contagious[3,5]. For example, one Ghanaian study found that a sizeable proportion of participants described epilepsy as a “repeated death” where sufferers die, revive, and die again, in a repetitive cycle[5].

The implementation of successful health care systems on the African continent to address the treatment gap for epilepsy will necessarily have to be accompanied by education campaigns to dispel the stigma surrounding epilepsy as a disease. Initiatives have already begun to be put in place by health authorities to address this social stigma, including a project launched by the International League Against Epilepsy (ILAE), International Bureau of Epilepsy (IBE), and the World Health Organisation (WHO) to “bring epilepsy out of the shadows”. There are also national-level projects currently being implemented across Africa to address such stigma[5].

Increased understanding and scientific education regarding the aetiology of epilepsy is, and will continue to be, a vital component of those education campaigns. An understanding of genetic epilepsy on the African continent is particularly important in this context, as the vast majority of epilepsies for which causes are unknown are likely to be genetic in origin[4]. A genetic explanation for the development of epilepsy that otherwise seems to be without cause may help dispel beliefs in supernatural causes and thus reduce fear and stigma.

1.2. Epilepsy and seizures: classifications and definitions

Epilepsy was defined in 2005 as a neurological disorder characterised by unprovoked, recurring seizures, the occurrences of which happen over a period of time greater than 24 hours[13]. In 2014, the ILAE refined that definition into a more practical and clinically useful form - epilepsy is thus currently defined as a disease in which an individual displays at least one of the following characteristics:

i. At least two unprovoked seizures occurring greater than 24 hours apart.

ii. One unprovoked seizure and a probability of further seizures similar to the general recurrence risk (at least 60%) after two unprovoked seizures, occurring over the next 10 years.
iii. Diagnosis of an epilepsy syndrome[14]. The position of the ILAE on the classification of epilepsy syndromes was published in 2017[15].

Once diagnosed, clinical management of epilepsy includes identification of its cause, administration of ASMs, and the treatment of co-morbidities[3]. In cases of paediatric epilepsy, psychoeducational interventions for primary caregivers are also a vital part of holistic treatment[16].

Epilepsy is a highly variable disease, which occurs in multiple conditions with many different causes, ages of onset, and manifestations. Classifications of epilepsy and epileptic seizures have changed over time. The ILAE's current revised classification of seizure type is operational, as opposed to being based on seizure mechanisms[17]. Seizures are classified as being focal, generalised, or of unknown onset, with subcategories of motor, non-motor, sustained awareness, or impaired awareness (Figure 3)[17].

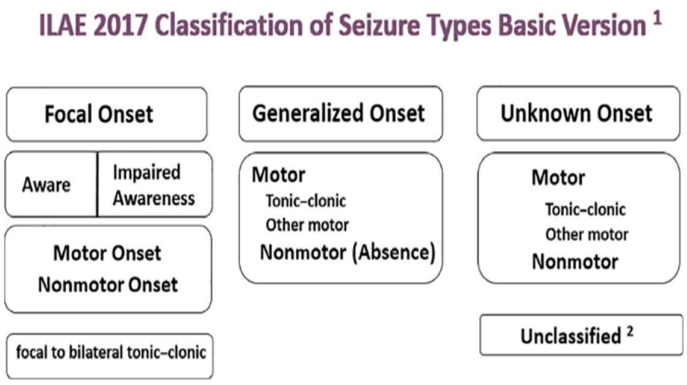


Figure 3 The ILAE classification of seizure types. (Figure adapted from Fisher et al. 2017)[17]

The current framework for the classification of epilepsies published by ILAE in 2017 takes into account seizure types, epilepsy types, epilepsy syndromes, aetiology, and co-morbidities (Figure 4)[15]. Classifications of seizures and epilepsy are an important diagnostic tool for practicing clinicians and are regularly re-evaluated and updated to reflect the growing body of knowledge and insight into the disease mechanisms and aetiology. Classification's primary

purpose is to aid patient diagnosis and treatment; it is also critical for epilepsy research, antiseizure drug development, and communication.

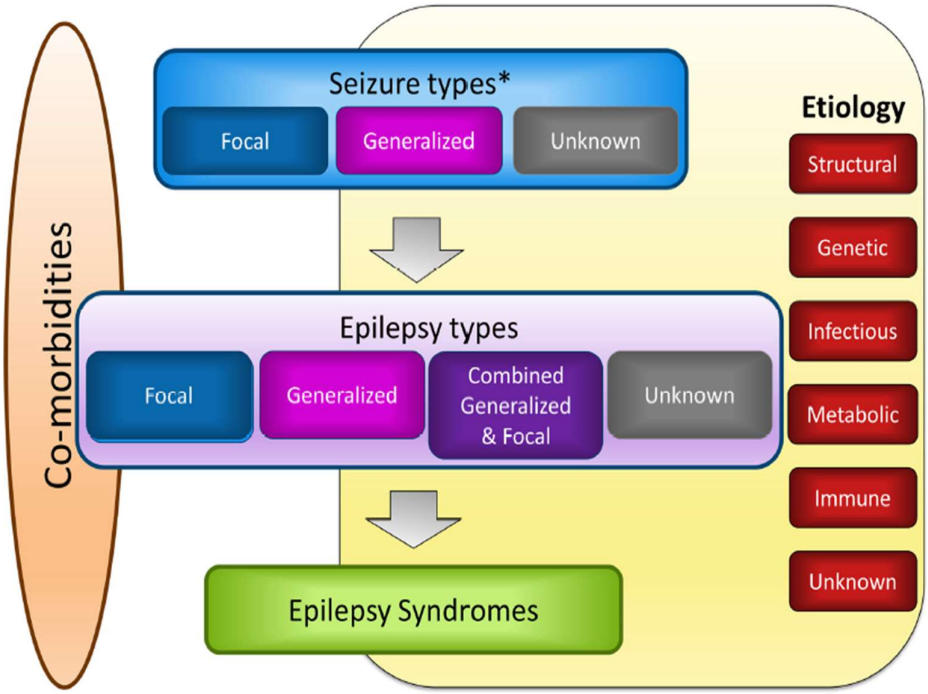


Figure 4 The ILAE framework for classification of epilepsies. (Figure adapted from Scheffer et al. 2017.)[15]

The highest incidence of epilepsy is observed during the period of infancy and early childhood, which is also when uncontrolled seizures have the most severe short- and long-term effects on the developing brain[18]. Paediatric epilepsy syndromes may be categorised as self-limited focal epilepsies, generalised epilepsies, or developmental epileptic encephalopathies[19]. Infantile and childhood-onset epilepsies, such as the DEEs, often have severe co-morbidities and neurodevelopmental challenges, and often display drug-resistance[20]. Drug-resistant, or refractory, epilepsy, is defined by the ILAE as failure to control seizures after treatment with two appropriate ASMs[21].

1.2.1. Developmental and Epileptic Encephalopathies

Childhood-onset epilepsies and epilepsy syndromes often occur with neurological co-morbidities. Multiple terms are used for neurological and neurodevelopmental phenotypes

which may include epileptic seizures. Of particular note are epileptic encephalopathy (EE), developmental encephalopathy (DE) (in which seizures may be absent), and DEE.

Epileptic encephalopathy refers to a condition where neurological, cognitive, and motor function deteriorate as a result of continued epileptic activity in the brain[22]. Developmental encephalopathy refers to a condition where an underlying genetic cause is associated with the observed neurological impairment, independent of epileptic seizures[23]. In contrast to both these conditions, DEEs are conditions where both an underlying genetic cause and epileptic activity are thought to contribute to neurological impairment[24,25].

Over 100 genes have been associated with the onset of DEEs, and DEEs are highly heterogenous[23]. Despite this heterogeneity, consistent characteristics across the DEEs include an early age-of-onset, resistance to pharmacological treatment, slowed intellectual development, or neuroregression, and multiple seizure types, among other electroclinical characteristics[23]. Conditions classified as DEEs include, among others, Ohtahara syndrome, early myoclonic encephalopathy, epileptic spasms syndrome, West syndrome, Dravet syndrome (DS) and Dravet-like syndromes, and epilepsy of infancy with migrating focal seizures (EIMFS)[23].

The acute severity of the DEEs, and their early onset, necessitates a rapid identification of effective treatment strategies to minimise damage to the developing brain. Genetic testing plays a crucial role in the application of precision medicine techniques for these patients, as identification of a specific disease-causing variants may direct the choice of pharmacological treatment (e.g., avoidance of sodium channel blockers in *SCN1A*-related DS) and prevent unnecessary diagnostic tests or inappropriate treatment interventions[23].

Severe infantile and childhood DEEs have been recognised as genetically heterogenous, and most commonly arising from dominant *de novo* pathogenic variants in affected individuals, although other inheritance patterns have also been observed[26]. The pathogenic variant rate in genetic diagnosis of DEEs is continually improving with ongoing epilepsy gene discovery, but additional, currently unknown mechanisms are also likely to play a role in the development of DEEs. Despite DEEs having the highest rate of actionable genetic findings, more than 50% of patients diagnosed with a DEE remain without a genetic diagnosis[27].

The identification of these mechanisms may be aided not only by further new gene discovery, but also may require an adjustment of diagnostic techniques – for example, DEE's may be caused by a high prevalence of post-zygotic somatic mutations, which may go undetected by current sequencing protocols, or may be due to the influence of epigenetic factors[27]. As advances in diagnostic techniques and further research enable greater success in identifying causative variants, the prognoses for patients diagnosed with DEEs are likely to improve, especially given the high rate of actionable genetic findings observed so far in DEEs.

1.3. Gene Discovery

Hundreds of genes have been described by researchers and clinicians as being causative of, or associated with epilepsy, with such genes categorised into five broad functional categories:

i. Genes involved in ion transport.
ii. Genes involved in cell growth and differentiation.
iii. Genes involved in the regulation of synaptic processes.
iv. Genes involved with the transport and metabolism of small molecules within and between cells.
v. Genes involved in the regulation of gene transcription and translation[28].

Many genes implicated in epilepsy are now routinely tested for in clinical settings, and gene discovery, enabled and enhanced by new laboratory techniques, continues to add to the list of causative or epilepsy-associated genes.

The majority of genes implicated in epilepsy encode ion channel proteins, enzymes, or enzyme modulators typically identified as important in the transmission of signals across neurological synapses[29,30]. However, the epilepsy-causing genes are not limited to those categories. For example, pathogenic variants in *WDR45,* a gene implicated in autophagy, have been observed as being causative for DEEs and other neurodevelopmental conditions[31]. The wide range of different neuronal functions (encoded by different genes) which, when impeded, may be causative of epilepsy is illustrated in Figure 5[32].

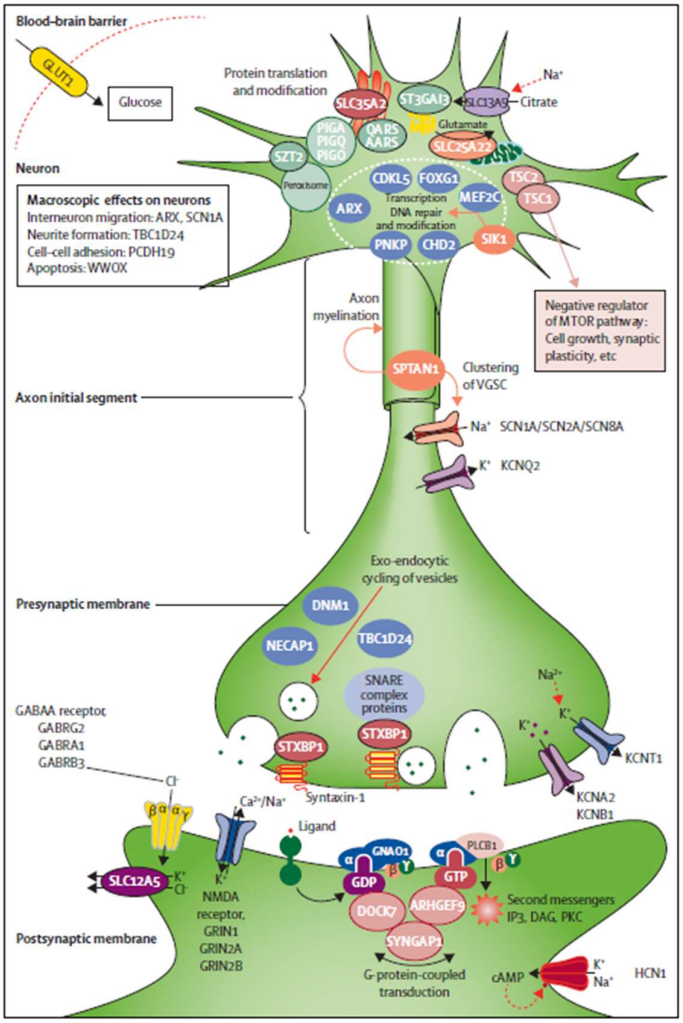


Figure 5 The disease mechanisms which may impede normal neuronal functioning and cause epilepsy in childhood. (Figure adapted from McTague, et al. 2016.)[32]

Equally important to gene discovery is the continuous review of candidate epilepsy genes through ongoing research and assessment of new evidence as it arises. For example, the *SCN9A* gene, which encodes the alpha subunit of a voltage-gated, type IX sodium channel, was previously thought to be a causative epilepsy gene. However, *SCN9A* has since been downgraded to modifier status, as subsequent research data did not sufficiently support disease-association[33]. Other genes previously thought to be implicated in epilepsy that have subsequently been shown to lack evidence to currently be considered disease genes, include *CLCN2, CACNA1H, EFHC1, GABRD,* and *SRPX2*[20].

As epilepsy is highly genetically heterogenous, focus on gene discovery for single-gene epilepsies will not address all genetic epilepsies, but identifying more commonly occurring single-gene epilepsies will aid in targeting resources, prioritisation of diagnostic testing, and the development of precision therapies[34]. Such novel gene discovery has rapidly accelerated through the use of techniques such as exome sequencing (ES) and genome sequencing (GS).

In addition to this accelerating gene discovery, a relatively recent approach to investigating the genetic aetiology of epilepsy is the investigation of the frequencies of variants in different genes in cohorts affected with epilepsy. Large scale sequencing studies have studied the enrichment of rare variants across multiple genes in individuals affected by epilepsy; one such study, published in 2019 by Heyne et al., analysed data from 6994 gene panels for neurodevelopmental disorders with epilepsy, and found the genes with the highest frequencies of ultra-rare variants in cases of neurodevelopment disorders and epilepsy to be *SCN1A*, *KCNQ2*, *SCN2A*, *CDKL5, SCN8A*, and *STXBP1*[35]. Another study, by Feng et al., found that, compared to controls, patients affected with epilepsy have an excess of ultra-rare deleterious variants in both constrained genes and genes associated with epilepsy[36]. In the context of DEEs specifically, DEEs have conventionally been considered monogenic disease, but recent research suggests that a substantial portion of DEE cases may be due to affected individuals carrying a large number of ultra-rare variants, resulting in an oligogenic mode of inheritance[37].

In addition to the study of rare variants, polygenic risk scores have been applied to the study of how combinations of common variants could contribute to the development of epilepsy[38]. This approach may be of particular use in elucidating the cause of generalised epilepsies, with current research suggesting the cumulative effect of common variants could explain up to a third of genetic risk for generalised epilepsies[38]. Large scale statistical approaches may also direct future gene discovery approaches - a 2020 study by Kaplanis et al. analysed sequencing data from 31 000 parent-offspring trios for gene-specific enrichment of *de novo* variants, identifying 28 novel genes associated with developmental encephalopathies, and predicting that as many as 1 000 causative genes are still to be described[39]. Studies such as these highlight the importance of a multifaceted approach to gene discovery, combining traditional candidate-gene research with newer statistical approaches which utilise large data sets.

1.4. Phenotypic Overlap and Genetic Heterogeneity

Clinical diagnosis of early-onset epilepsy in children can be accompanied by frequent co-morbidities, such as autism spectrum disorder (ASD) and a range of behavioural and movement disorders. Epilepsy may also occur in children affected with other syndromic genetic disorders, such as Angelman syndrome[40]. Genetics-based diagnosis, as well as electro-clinical diagnosis, thus plays an increasingly important role in the aetiological diagnosis of epilepsy and selection of targeted treatments, as knowledge of the roles various genes play in the development of epilepsy increases. Aside from the clinical benefits of a definitive diagnosis, insight into the genetic diversity of epilepsy also expands the phenotypic spectrum of features associated with certain genes, and continuously demonstrates the high genetic heterogeneity of the paediatric epilepsies.

Broad phenotypic spectrums are often observed in DEEs, where epilepsy phenotypes may include broader neurological features. For example, *STXBP1* plays a role in neurotransmitter secretion and *de novo* pathogenic variants in this gene are among the most frequent causes of both epilepsy and other neurological co-morbidities; patients with epilepsy caused by *STXBP1* variants also demonstrate intellectual disability[41]. However, researchers have found little correlation between the severity of intellectual disability and seizure severity, suggesting that seizures and intellectual disability, in those cases, are two independent phenotypes arising from the variant enabling more than one pathophysiological mechanism in the brain[41]. (Thus, by definition, *STXBP1* epilepsies are one facet of a complex neurodevelopmental disorder, a DEE, rather than a pure EE[41].)

Causative variants in *STXBP1* have been observed in cases of Ohtahara syndrome, West syndrome, non-syndromic early onset epilepsy and encephalopathy, Dravet syndrome, non-syndromic epilepsy and intellectual disability, and ASD, demonstrating a wide phenotypic spectrum[41]. Cases of *STXBP1* epilepsy also highlight the difficulty in reconciling electro-clinical syndromes to distinct genetic causes, particularly when those genetic causes may be interacting with genetic modifiers and environmental factors to produce the eventual phenotype[41].

The genetic heterogeneity and phenotypic overlap of genetic epilepsies complicates the process of choosing testing approaches based on genotype-phenotype correlations. Multiple genes, including *SCN1A, SCN2A, SCN8A, KCNQ2, SLC2A1, GABRA1, GABRG2, GRIN2A,* and *PRRT2* have been implicated in both mild and severe cases of epilepsy[42]. Additionally, genetic pleiotropy has also been observed in studies of epilepsy; for example, variants in a single gene, *SCN2A,* have been associated with as wide a range of diagnoses as mild self-limited familial neonatal epilepsy, Ohtahara syndrome, epilepsy of infancy with migrating focal seizures, West syndrome, Lennox-Gastaut syndrome, and severe, unclassifiable DEEs[23]. Even variants of the same functional characteristics in the same gene may demonstrate phenotypic range – for example, gain-of-function (GOF) variants in *CACNA1I* have been implicated in a spectrum of phenotypes ranging from borderline intellectual functioning to severe DEEs[43]. Even the same precise variant may demonstrate phenotypic range across different individuals. In contrast, research has shown that pathogenic variants at different loci may result in similar phenotypes. To date, benign familial epilepsy syndromes have been associated with at least 10 genes, Ohtahara syndrome with variants in at least 15 genes, and over 45 genes have been implicated in infantile spasms syndromes, including West syndrome[23].

Phenotypic overlap thus presents challenges to predicting possible genetic causes of epilepsy based on electroclinical diagnoses. Despite these challenges, increasingly common genetic testing and larger scale studies are continuously elucidating further insights into such phenotypic overlap. The development of standardised language and phenotypic ontologies has also allowed for the increasingly common application of quantitative phenotypic analyses. The Human Phenotype Ontology (HPO) system, used in the study of multiple different diseases and disorders, is increasingly being used in the study of epilepsy. The use of the HPO system in the context of epilepsy has resulted in an expansion of the number of descriptive concepts which may be applied in description of clinical phenotypes, including seizure descriptions[44].

A 2021 study by Crawford et al. found that the use of HPO terms in the analysis of phenotypic data of *SCN2A* epilepsies helped delineate correlations between the functional characteristics of a variant and particular phenotypes – such as the association of protein-truncating variants with ASD and behavioural abnormalities, and of missense variants with neonatal onset epileptic spasms[45]. The results of the study included a broad categorisation of *SCN2A* clinical presentations, typically thought to have considerable phenotypic overlap, into three

categories: benign familial infantile seizures, ASD, and DEE[45]. The application of ontologies and systematic analysis to elucidate phenotypic overlap thus seems promising. Such standardised terminology would also aid in the collaboration across multiple facilities and research groups[46].

Adding to the genetic heterogeneity in epilepsy is the possible influence of modifier genes, which has not been fully elucidated, but may play a role in variant penetrance and effect on phenotype[47]. A 2018 study by Bagnasco et al. described a case in which a particular variant observed in *SCN8A* (c.3943G>A) that had been previously identified as causative in two cases of DEE was found to be causative in a mild case of what was described as benign focal epilepsy with normal cognitive development[48]. The authors hypothesise that the discrepancies between the effects of this variant could be due to modifier effects, additional variants in different genes, or mosaicism. Modifier effects may be further complicated if several modifier genes act in combination to affect disease risk and presentation, rather than acting individually[49].

Mosaicism, found to occur relatively frequently in many epilepsy genes, including *CDKL5, GABRA1, GABRG2, GRIN2B, KCNQ2, MECP2, PCHD19,* and *SCN2A*, is another factor which may affect disease severity and recurrence risk in families (if occurring in parents)[50]. For example, understanding the risk of recurrence in DS cases may be complicated by parental mosaicism, with studies observing low grade mosaicism in 7-13% of families with an affected child[51]. Cases of mosaicism in affected individuals may also present challenges to testing and variant identification during routine diagnostic processes. In most cases where the causative variant is *de novo*, the variant arises during gametogenesis in the parent, and thus the variant would be ubiquitously present and detectable in the tissues which are routinely used to collect DNA for testing, such as blood lymphocytes and buccal cells. In contrast, post-zygotic mutations, resulting in mosaicism, may result in a variant being present in particular tissue not used during routine testing, and which may be difficult to access when performing routine testing on a patient[46].

1.5. Inheritance patterns

Ongoing research and epilepsy gene discovery has highlighted the multiple types of inheritance patterns that contribute to the development of epilepsy, ranging from monogenic to complex inheritance, resulting from combinations of multiple genetic and non-genetic factors[52]. Complex inheritance of epilepsy may include: oligogenic inheritance, where a small number of variants combine to induce disease; polygenic inheritance, where many variants (potentially thousands) contribute to the development of disease; and omnigenic inheritance, where genes not obviously connected to a disease may contribute to disease development through interacting gene networks[38]. Most cases of epilepsy, especially the DEEs, are monogenic and caused by *de novo* variants in autosomal dominant (AD) genes. Autosomal recessive (AR) and X-Linked or mitochondrial genes have also been implicated, though to a lesser degree.

A particularly interesting example of different factors combining to produce a counter-intuitive inheritance pattern is *PCDH19*-associated epilepsy. The epilepsy syndrome associated with variants in *PCDH19* (located on the X-chromosome) displays what is referred to as a male-sparing pattern of inheritance – heterozygous females are affected while hemizygous males are "spared" the development of disease. This is counterintuitive, as most diseases caused by genes carried on the X-chromosome manifest to a greater extent in hemizygous males, while heterozygous females are often apparently unaffected carriers (such as in haemophilia). Research by Pederick et al. has shown that the cellular adhesion properties of the protein encoded by *PCDH19* are likely to be responsible for this atypical pattern of inheritance. Mosaic expression of *PCDH19* (the presence of heterozygosity for a pathogenic variant) was observed in mouse models to lead to abnormal cell sorting during neurological development, which correlates with the cortical malformation observed in individuals affected with *PCDH19*-associated epilepsy[53]. The presence of a heterozygous genotype resulting in pathological cell sorting thus explains why homozygous females and hemizygous males are unaffected while heterozygous females are affected[53]. The complicated disease mechanism and relatively unusual inheritance pattern of *PCDH19*-associated epilepsy is an example of how understanding of genetic epilepsies can be complicated, and counter-intuitive manifestations of disease require collaboration across many biomedical disciplines to fully elucidate the mechanisms of disease and aid in the treatment of patients.

1.6. Genotype-phenotype correlations

As gene discovery accelerates and molecular genetic testing becomes part of standard practice in clinical epileptology, new and useful insights into genotype-phenotype correlations are being revealed. The ClinGen Epilepsy Gene Curation Expert Panel is an undertaking currently underway utilising the skills of experienced neurologists to continuously review the validity of newly discovered genotype-phenotype correlations[54]. This is an example of how, as genetic testing accelerates, a concurrent field of research into gene-disease relationships will be necessary to inform future decisions about gene selection for diagnostic tests and maximise the benefit of precision medicine techniques. Certain causative genes and variants are known to have a characteristic clinical presentation, such as pathogenic *SCN1A* variants causing DS, which can help direct and streamline the genetic testing and diagnostic process[55]. Other examples include *STXBP1*-associated DEE, which usually presents with movement disorders such as hand stereotypies, and *PCDH19*-associated DEEs which feature clustered focal seizures.[4,56].

One example of a subtype of epilepsy for which a genetic basis is well characterised is Dravet syndrome. In cases of suspected DS, screening for causative variants in *SCN1A* is routine, as *de novo* pathogenic variants found within *SCN1A* have been associated with 80% of DS cases[57]. In cases of DS, consolidation of the diagnosis through genetic analysis goes beyond genetic counselling as it informs treatment (for example, the avoidance of ion channel blocker ASMs, the use of which exacerbates the disease mechanism and thus increases the severity of symptoms) and can direct specific interventions, such as facilitating access to stiripentol in regions where such access is limited (e.g., the South African state health service) and the introduction of a ketogenic diet[58]. As the example of DS demonstrates, aside from value in making diagnoses, the specific identification of pathogenic variants and their associated disease mechanisms will also enable better understanding of the neuropathophysiology of epilepsy and the mechanisms by which it develops, and thus lay the groundwork for the development of targeted therapies[26].

However, it is important to note that the relative genetic homogeneity of DS is unusual amongst genetic epilepsies, and genetic epilepsies demonstrate substantial genetic heterogeneity, as described above[57]. Even in cases of strong genotype-phenotype correlations, genetic heterogeneity and pleiotropy must be considered. For example, *SCN1A*

is implicated in the development of a range of epilepsy phenotypes (and non-epilepsy phenotypes) other than DS, including, amongst others, genetic epilepsy with febrile seizures plus, myoclonic-atonic epilepsy, EIMFS, early onset *SCN1A* DEE, hemiplegic migraine, and ASD[59,60]. Cases of DS have also been observed to occur in the absence of pathogenic variants in *SCN1A* – causative variants in DS patients have been observed in *SCN1B, SCN2A, SCN8A, SCN9A, PCDH19, GABRA1, GABRG2, STXBP1, HCN1, CHD2,* and *KCNA2*[61–64].

1.7. Precision Medicine Approaches to the Treatment of Epilepsy

Precision medicine, which aims to treat every individual on the basis of their genetic composition and personal environmental factors, both within and outside the context of rare diseases, has increasingly been identified as a powerful tool in enabling the successful treatment of epilepsy across the globe. Precision medicine is not a new concept and is not unique to the field of genetics, but its applications have been boosted by genetics to a considerable degree[65]. Additionally, precision medicine should be considered a series of approaches to treatment working in tandem, rather than a singular treatment in isolation[65].

Precision medicine approaches include, amongst others:

i. biometric monitoring (which formed the basis of this study's parent study, *Precision Management of Epilepsy*);
ii. genetic testing approaches which allow for the identification of disease-causing variants in affected individuals, and thus the allow for treatments specific to those disease-causing variants, should they exist; and
iii. pharmacogenomic investigations to improve the efficiency and efficacy of prescribed medication.

The latter two concepts are discussed in the context of epilepsy in the following sections.

1.7.1. Efficacious Genetic Testing in Epilepsy: NGS Approaches for Genetic Diagnoses

Current evidence suggests that genetic testing for neonatal- and infantile-onset epilepsies produces the highest diagnostic yield and most informative and actionable results[66]. Next-

generation sequencing (NGS) approaches are a powerful way of approaching the challenges of genetic heterogeneity found within epilepsy syndromes, the phenotypic heterogeneity associated with different mutations in the same gene, and the possibility of mosaicism[50,67]. Additionally, NGS-based genetic testing allows for the diagnosis of epileptic disorders which do not demonstrate typical clinical or electrographic features (such as many metabolic epilepsies) and can allow for a reliable diagnosis without the need for invasive procedures such as biopsies, metabolic testing, or cerebrospinal fluid analysis[68]. The ability to circumnavigate invasive, often painful diagnostic procedures is of particular benefit in paediatrics, where such procedures are often traumatic for both children and their families.

By 2017, approximately one thousand genes had been associated with epilepsy, and gene discovery for epileptic disorders has continued at a substantial rate[29]. Next-generation sequencing is thus the most efficient and cost-effective approach to genetics-based diagnosis of epilepsy, due to its ability to sequence multiple genes simultaneously, and is the routine approach to the diagnosis of genetic epilepsy in HICs[4]. In some studies, the use of NGS approaches has increased the likelihood of the identification of a causal variant in epileptic encephalopathy cases from less than 10% to more than 30%[69]. Phenotypic features which are associated with a higher diagnostic yield include early onset of seizures, drug resistance, and developmental co-morbidity[70–72]. Additionally, diagnostic yield has been observed to differ by testing method, with GS, ES, and gene panels producing diagnostic yield of up to 48%, 45%, and 25%, respectively[46]. The genes most commonly identified as having causative variants for epilepsy in NGS studies are *SCN1A, KCNQ2, CDKL5, SCN2A*, and *STXBP1*, and monogenic cases remain the main target of clinical testing[46,70]. Testing by chromosomal microarray has been observed to produce a diagnostic yield of 5-15% in cases of DEE[46].

The most commonly used approaches in the clinical setting for epilepsy are ES, which involves the sequencing of all the protein coding regions (exons) in the entire genome, and gene panels, which sequence a finite list of genes selected for the panel based on the disorder under investigation. Genome sequencing is increasingly used in the diagnostic laboratories of HICs, though this is still prohibitively expensive in LMICs[46]. All these approaches allow for an increase in diagnostic efficiency that have led to substantial improvements on previous diagnostic rates for epilepsy, with the use of traditional direct cycle sequencing for the genetic diagnosis of epilepsy essentially rendered obsolete[46].

The NGS approach most commonly taken in the diagnostic setting is the use of gene panels. Gene panels can give higher average coverage than exomes, resulting in better depth of coverage and greater power to identify small deletions and mosaicism, and are additionally less costly and technically easier to run, with smaller data storage requirements[69]. Panels also avoid the ethical challenges that may be presented by incidental findings, due to their specificity. The broader the diagnostic approach, the more likely the chance of an incidental finding occurring, though only actionable incidental findings would be reported. The yield of actionable incidental findings in cases of epilepsy ranges from 2-4%[46]. However, a challenge of the gene panel approach is the process of choosing what genes to sequence, in order to maximise diagnostic yield. In a review of commercially available gene panels, Chambers et al. found gene panels for epilepsy to range in size from 70 to 465 genes, with an overlap of genes implicated in many different diseases[73].

There are multiple factors influencing the diagnostic yield of panels, and the size of a panel does not necessarily directly correlate to diagnostic yield. Diagnostic yield tends to be highly variable across different panels, with a range of 10% to 72% reported in across different studies[74]. This is likely due to different study designs and particularly the participant selection process, including the level of phenotyping performed, and the inclusion and exclusion criteria. Though ES and GS do have the highest diagnostic yields, the difference in yield between ES and a well-designed gene panel may be negligible if the targeted population sequenced on the panel is appropriate for that panel and selected based on robust phenotyping. This demonstrates that, despite its powerful applications, NGS remains a tool to answer a biological question, and the context in which NGS technology is applied has a greater effect on the results than the choice of NGS technology itself.

There is a substantial research effort in the field of epilepsy genetics to maximise the diagnostic yield of gene panels. The main approach involves gathering robust statistical evidence for the inclusion of genes on panels, as well as population data to potentially design population-specific gene panels[75]. In the case of epilepsy, population specific panels are likely to be relevant mainly in the case of recessively inherited forms of the disease, as the frequencies of *de novo* mutations in dominantly inherited epilepsy-associated genes, which make up the majority of causative variants, are likely to hold steady across different populations. A demonstration of this is apparent in that two studies, one conducted in a Scottish cohort and the other conducted in a Chinese cohort, both found *PRRT2* to be the most commonly reported causative gene in paediatric epilepsy[75,76]. (As yet another example

of genetic and phenotypic heterogeneity, *PRRT2* has also been implicated as a causative gene in the movement disorder, paroxysmal kinesigenic dyskinesia[77].)

Despite the widespread use of gene panels, ES and GS play an important role in the diagnosis of epilepsy, especially in cases where the gene panel approach has not produced a result. Despite being more technically challenging and more expensive, these approaches, ES in particular, are gaining prominence as the first-tier genetic test for epilepsy in the HICs[46]. A particular strength of ES and GS approaches is the broadening of the phenotypic spectrum for both genes and variants and the discovery of modifier and/or additional variants. For example, an exome-based study by Allen et al. broadened the phenotypic spectrum for variants in *KCNB1*, and identified possible modifier variants in *CPA6*[78]. An additional strength of ES and GS approaches is that they allow for re-analysis of patient genomic data as new knowledge becomes available and new genes are identified as causative of or associated with epilepsy. Genome sequencing may also be used for genomic copy number variant (CNV) detection, circumnavigating the need for microarray analysis. This aspect of GS may be found to be increasingly important considering the accumulating data considering the prevalence and importance of CNVs in genetic epilepsies[79,80]. Examples of epilepsies occurring due to CNVs include 22q11.2 deletion syndrome, 15q13.3 microdeletion-associated epilepsy, and epilepsies occurring with co-morbidities of ASD, intellectual disability, and schizophrenia[37].

NGS-based testing remains efficacious even in epilepsies which may be diagnosed by non-genetic means, such as metabolic epilepsies. Examples include pyridoxine-dependent epilepsy or GLUT1 deficiency syndrome, where genetic testing may assist in diagnosing atypically presenting affected individuals, or individuals who have difficulty accessing biochemical diagnostic tests[81]. Efficient means to diagnose metabolic epilepsies have been recognised as increasingly important, as improvements in diagnostics have demonstrated a metabolic basis for a substantial portion of recurring seizures in neonates and children[82]. Such metabolic disorders include inherited metabolic disorders with predominantly epileptic manifestations and metabolic disorders where epilepsy occurs as part of an overarching neurological phenotype[82].

Presently, the two main international market leaders in NGS technologies are Illumina (San Diego, California, USA) and Thermo Fisher Scientific (Waltham, Massachusetts, USA). Illumina sequencing technology uses proprietary fluorescently labelled, reversible terminator

nucleotides in order to perform sequencing by synthesis, reading the sequence via the fluorescent signals given off by nucleotides incorporated into the DNA. The Ion Torrent platform from Thermo Fisher Scientific also performs sequencing by synthesis, but uses semiconductor sequencing technology to measure the release of hydrogen ions as nucleotides are incorporated into DNA and the hydrogen bonding between complementary nucleotides occurs. The change in concentration of hydrogen ions in the solution is detected as changes in pH.

1.7.2. Clinical Utility of Genetic Diagnoses: Variant Interpretation and Precision Medicine Approaches

The identification of disease-causing variants in sequencing results is often challenged by the common co-morbidity of epilepsy with other neurodevelopmental disorders which often have a similar or shared genetic basis (such as ASD and intellectual disability)[83]; complicated inheritance patterns, such as incomplete penetrance or mosaicism; unexplained phenotypic heterogeneity; and lack of knowledge as to what additional factors may influence phenotypic expression of pathogenic variants[20]**.** All these factors contribute to the complexity of the process of filtering through variants identified by sequencing and assigning pathogenicity classifications to those variants. Furthermore, the ability to assess variant pathogenicity and functionality has not kept pace with the rapidly increased rate of gene discovery (and variant detection), potentially complicating the process of diagnosis and counselling of affected families[84]. Genetic counselling may be particularly difficult in cases where variants are classified as variants of uncertain significance (VUS). However, cases involving a VUS may be resolved at a later date upon reanalysis of the variant in light of new research and new data. Any reanalysis of genomic data should also be accompanied by up-to-date phenotyping of the patient in question[46].

Despite the challenges described above, the importance of a genetic diagnosis cannot be overstated, as a genetic diagnosis can help direct treatment, reduce the need for additional investigations, predict future health outcomes, and can have immense psychological benefit for affected families, such as alleviating feelings of culpability and guilt by giving them diagnostic closure[26]. Genetics-based diagnosis is therefore an important part of the process of multidisciplinary treatment, and despite the complexity of epilepsy and the challenges of diagnosis and treatment, several precision medicine approaches have already been

implemented for the treatment of epilepsy. Those approaches include informed choices regarding ASM use, the repurposing of drugs approved for the treatment of other diseases, the use of enzyme replacement therapy, and the development of new gene therapy approaches[37]. An example of drug repurposing is the approval of fenfluramine, previously used as an appetite suppressant, for the treatment of DS in 2020[49]. Research into new gene therapy approaches for the treatment of epilepsy, such as the use of anti-sense oligonucleotides (ASOs), currently involves the use of animal models and preparation for human clinical trials[37,85].

The use of NGS testing to enable genetic diagnoses and identify specific pathogenic variants is of great importance for the use and development of targeted therapies. Genetic heterogeneity in the epilepsies extends beyond the influence of multiple genes to heterogeneity of different variants in the same gene. The functional nature of causative variants, whether they are loss-of-function (LOF) or GOF, has been observed to influence disease pathophysiology. The precise nature of pathogenic variants is especially relevant to treatment decisions, as epilepsy caused by a LOF variants may have to be treated differently to epilepsy caused by GOF variants in the same gene[47]. The *KCNQ2* gene provides an example of this phenomena, demonstrating that knowledge of the precise causative variant is essential, and assumptions about treatment based on the gene rather than the causative variant may be inaccurate[86]. Another example is the gene *SCN2A,* found to be causative of a range of epilepsies – GOF variants in *SCN2A* may be treated effectively with sodium channel blockers, whereas in cases of LOF variants, sodium channels blockers must be avoided[23].

Current therapies for many epilepsies remain imprecise and focused on empirical seizure reduction rather than targeting specific disease mechanisms, but there are several existing disease-mechanism specific approaches for epilepsy, and identifying a causative gene can help direct ASM prescription in up to 76% of cases of paediatric epilepsy[46,49]. Classical examples of disease mechanism-specific approaches to treating epilepsy include: treatments for pyridoxine-dependent epilepsy cause by *ALDH7A1* variants, treatment of pyridoxamine-5'-phosphate oxidase deficiency, treatment of pyridoxal-phosphate homeostasis protein deficiency, treatment of GLUT1 deficiency, the treatment of *PRRT2*-related epilepsy with carbamazepine, the treatment of *KCNT1* epilepsy with quinidine, the use of memantine, ketamine, and serine in the treatment of *GRIN1A, GRIN2A, GRIN2B,* and *GRIN2D* related epilepsies, and the use of ASMs in *SCN1A*-related epilepsies directed by whether variants

demonstrate GOF or LOF disease mechanisms[87]. Further examples of genetically directed therapies are given in Figure 6.

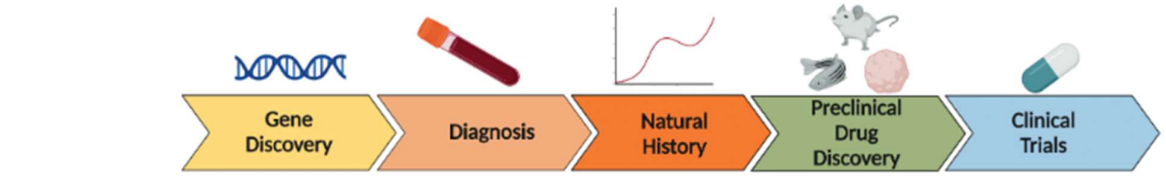
As precision medicine approaches are continuously developed and more widely implemented, additional challenges are likely to arise which would need to be addressed. Precision medicine approaches are not always easily accessible to all patients who would benefit from them, and additionally, may actually not work successfully[65]. A successful application of precision medicine would entail the treatment of a well-defined genetic mechanism, considering patient-specific factors, resulting in complete seizure control and freedom from co-morbidities[49]. Failure to implement such a successful application may be due to numerous factors, such as the mechanism(s) generating the clinical phenotype being inadequately understood, and the influence of additional genetic and environmental factors upon individual patients[65].

The development of a comprehensive framework for evaluating the success and scalability of precision medicine approaches will become increasingly important as precision medicine strategies are more frequently applied and new treatments are developed. Investigators propose that the development of new precision medicine strategies would ideally include phases of gene discovery, diagnosis, characterisation of natural histories, preclinical drug discovery, and clinical trials[49]. Each of these phases has its own associated challenges, as well as notable successes which may be built upon to allow for continued progress in the field and the continual improvement of clinical outcomes in patients (Figure 7).

Genes	Proteins	Main pathophysiology	Potential precision treatment approaches	Evidence	Reference
ALDH7A1 *PNPO* *PROSC*	Aldehyde-dehydrogenase Pyridoxine phosphate oxidase Pyridoxine phosphate binding protein	Vitamin B6 deficiency	Supplementation with pyridoxine Supplementation with pyridoxal-5-phosphate	+	Mills *et al.*, 2014 [69] Darin *et al.*, 2016 [70]
CAD	Trifunctional protein (CPSase, ATCase, DHOase) in pyrimidine biosynthesis	Deficiency in pyrimidine biosynthesis	Supplementation with uridine	+	Koch *et al.*, 2017 [71]
CHRNA4 *CHRNB2* *CHRNA2*	Nicotinic acetylcholine receptor (AChR)	Desensitization of the nicotinic AChR	Nicotine	+	Fox *et al.*, 2021 [72] Lossius *et al.*, 2020 [73]
GRIN1 *GRIN2A* *GRIN2B* *GRIN2D*	Glutamate receptor (NMDAR)	GoF/LoF	Memantine, dextrometorphane, ketamine for GOF, Serine for LOF	+	Pierson *et al.*, 2014 [74] Gale *et al.*, 2021 [75] Amador *et al.*, 2020 [76] Soto *et al.*, 2019 [63] Krey *et al.*, 2022 [64]
KCNA2	Voltage-gated K^+ channel $K_V1.2$ (A-type)	Loss or gain of function (or a mixture of both)	4-aminopyridine for GOF or some GOF+LOF variants to reduce channel overactivity	+	Syrbe *et al.*, 2015 [77] Hedrich *et al.*, 2021 [78]
KCNQ2 *KCNQ3*	Voltage-gated K^+ channels $K_V7.2$, $K_V7.3$ (M-type)	Loss or gain of function, depending on variant	Na^+ channel blockers for LOF variants (indirect effect blocking increased neuronal firing induced by reduced activity of K^+ channels); $K_V7.2/K_V7.3$ channel activators, such as ezogabine/retigabine	+	Pisano *et al.*, 2015 [79] Sands *et al.*, 2016 [80] Nissenkorn *et al.*, 2021 [81] Orhan *et al.*, 2014 [82] Millichap *et al.*, 2016 [83] Vanoye *et al.*, 2022 [84]
SCN1A *SCN2A* *SCN8A*	Voltage-gated Na^+ channels $Na_V1.1$, $Na_V1.2$, $Na_V1.6$	LOF or GOF of Na^+ channel function depending on individual variants	Na^+ channel blockers for GOF (to reduce channel overactivity), avoid such drugs for LOF variants (which may enhance reduced channel activity)	+	Guerrini *et al.*, 1998 [85] Wolff *et al.*, 2017 [86] Johannesen *et al.*, 2021 [43]
SLC2A1	Glucose transporter type 1 (GLUT1)	Reduced glucose transport across the blood-brain barrier	Ketogenic diet, providing ketone bodies as alternative fuel instead of glucose	+	Klepper *et al.*, 2020 [87]
TSC1 *TSC2*	Hamartin, Tuberin	mTOR disinhibition	Everolimus, Sirolimus (mTOR inhibitor)	++	French *et al.*, 2016 [88]

+: evidence from retrospective case series or clinical experiences from study groups; ++: evidence from a controlled clinical trial; GoF: gain of function; LoF: loss of function; AChR: acetylcholine receptor; AMPAR: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor; mTOR : mechanistic target of rapamycin.

Figure 6 Examples of precision medicine treatments for the treatment of epilepsy caused by specific genes (Figure adapted from Krey et al. 2022)[42].



	Gene Discovery	Diagnosis	Natural History	Preclinical Drug Discovery	Clinical Trials
Key Accomplishments	International efforts (e.g. Epi4K, EuroEPINOMICS, Epi25K): new genes, new mechanisms	Increased accessibility & speed of commercial genetic testing	Natural histories provided nuanced understanding of rare diseases (e.g. *cdkl5*, *STXBP1*)	• New preclinical models • Emerging use of ASO & AAV-based therapies	• Success of drug repurposing • Personalized ASO proof of concept (Milasen)
Key Challenges	Complexity of genotype-phenotype correlation	Unequal access to genetic testing and genetic counseling	Genetic discovery is outpacing clinical knowledge of pathogenesis and disease course	• Further validation and application of preclinical models • Infrastructure & common standards for gene-based therapies	• Traditional randomized controlled trials not feasible with some rare diseases • Need for non-seizure outcome measures & scales to measure clinically relevant changes • Equal access
Strategies for continued progress	• Data sharing & transparency • Innovation in bioinformatics • Natural history studies • Further understanding of polygenic risk, roles of intronic DNA	• Demonstrate cost effectiveness of genetic testing • Advocacy to shift standard of care • Increased neurogenomics training for providers • Increased patient education • Telemedicine	• Virtual hubs to efficiently coordinate multiple streams of information • Large-scale data harmonization • Multi-center collaboration	• Standardize production, monitoring & ethical oversight of gene-based therapies • Greater collaboration between physicians, scientists, patient advocates	• Common standards for "n of 1" or small trials to enhance rigor & generalizability • Patient-centered outcomes • Collaboration with patient advocacy groups • Leverage natural history for trial design

Figure created with Biorender.

Abbreviations: AAV, adeno-associated virus; ASO, antisense oligonucleotide.

Figure 7 Stages in the development of a precision medicine treatment, and their associated accomplishments, challenges, and future strategies (Figure adapted from Knowles et al. 2022)[49].

1.7.3. Pharmacogenetics and the Treatment of Epilepsy

The treatment of epilepsy involves use of ASMs, which are selected for each patient on the basis of his or her electroclinical presentation, age, and other clinical features. Conventional ASMs are used to prevent seizures, but do not treat the underlying cause of the epilepsy[88]. Drug resistant epilepsy (also referred to as pharmacoresistant or refractory epilepsy) is estimated to affect up to 30% of all epilepsy patients and is defined by the ILAE as failure to control seizures after treatment with two appropriate ASMs[89]. Drug resistance and ongoing seizure activity is a direct cause or contributor to the many clinical sequalae and overall disease burden associated with epilepsy. Individuals who are resistant to one ASM are more likely to be resistant to other ASMs, despite different mechanisms of action[90]. Even within

cases of successful treatment, there is a wide variety of responses to ASMs, including adverse drug reactions (ADRs)[91]. Efforts to improve the efficacy of pharmacological treatment for epilepsy include trying different combinations of drug regimens and the development of new ASMs, such as brivaracetam, which may be used for adjunctive treatment in combination with more traditional ASMs[92].

Currently, about 30 ASMs are available for symptomatic treatment of epilepsy, but the development of new treatments to address pharmacoresistance is limited in utility due to the diversity of epilepsy disorders and seizure types, and the complex patterns of drug resistance observed in patients with refractory epilepsy[90]. Observed patterns of drug resistance include:

i. *De novo* resistance where the patient never enters a period of seizure freedom while being treated with the ASM in question.
ii. Delayed drug resistance where the patient initially becomes seizure free, but seizures subsequently reoccur and become uncontrollable.
iii. Fluctuating resistance where seizures alternate between being controlled and uncontrolled.
iv. Cases in which the epilepsy is initially drug resistant but over time responds to treatment[90].

As described above, there are substantial challenges in addressing pharmacoresistance in epilepsy by developing and using new drugs. The efficiency of ASM prescription may, however, be improved through the application of pharmacogenetics. Pharmacogenetics is the study of what genomic variants may influence an individual's metabolism of, and thus response to, a drug, and the mechanisms of how that occurs. In addition to improving the efficiency of prescription, pharmacogenetics can also aid in the characterisation of different types of drug resistance. This may be done by using pharmacogenomic investigations to identify if a patient's drug resistance is arising from poor metabolism of an ASM, potentially rectifiable by a dosage adjustment, or if that ASM is simply not effective against the pathophysiology of the disorder treated. Adverse drug reactions are also a significant problem associated with ASMs, and also are a focus of pharmacogenomic research[93].

However, the clinical utility of pharmacogenetics in epilepsy remains limited by the lack of data available in the field[94]. Variation in the polymorphic cytochrome P450 (*CYP*) genes, such as *CYP2C9* and *CYP2C19*, has been found to influence ASM metabolism, particularly the

metabolism of phenytoin, but the evidence for *CYP* genes broadly influencing ASM metabolism is very limited, and much research in the field requires validation[90]. The role *CYP* genes may play in ASM metabolism are particularly poorly characterised in Africa[91]. Additionally, there is much greater *CYP* variation in African populations than in other populations, which may contribute to ASM responses and ADRs occurring differently in African populations than what would be expected based on studies of other populations[95].

It is thus necessary for better characterisation of African pharmacogenomic data, particularly in regard to ASMs, as effective and efficient treatment regimens tailored to an African population could contribute positively to reducing the high burden of epilepsy on the continent. In the context of the resource constraints in Africa, investigating the pharmacogenetics of the most widely used and financially accessible ASMs (such as phenobarbital, which is the most commonly prescribed ASM worldwide due to its low cost) should be prioritised[3]. More efficient prescription of even a few widely used ASMs could go a long way towards closing the treatment gap in SSA countries.

1.8. Precision Medicine for Epilepsy in Africa: Research and Translation

The majority of epilepsy research performed in Africa focuses on recognised aetiology, epidemiology, and management of the disease[4]. Although the implementation of wide scale genetic research in epilepsy in Africa could be challenging, such research could be extremely informative as well as beneficial on multiple levels. The potential benefit could extend beyond Africa and benefit populations worldwide in diagnosing and treating genetic epilepsies as African populations have the highest levels of genetic diversity of all human populations. It is also important to ensure genetic research becomes translatable in an African context and improves local skills development – such research should thus ideally be performed in Africa and include the creation of population specific databases to enable the translation of new scientific advances to clinical practice[4]. From the viewpoint of holistic health, it is imperative that concerted efforts are made to address the discrepancy in genetic diagnosis and research between Africa and the rest of the world. The implementation, over time, of genetic-based diagnostics for diseases such as epilepsy, would contribute to reducing that discrepancy.

Typically, research is carried out prior to the implementation of genetics-based diagnostic techniques, but for epilepsy in SSA, implementation of diagnostic testing in the clinical sector as part of a translational research project could be possible due to the high incidence of *de novo* variants in genetic epilepsies, particularly severe paediatric epilepsies. In this way, translation of research performed in other regions can be extended more broadly in the African context[6] .

At present there are no available NGS-based genetic tests for paediatric epilepsy in the South African public healthcare sector, although testing is available if samples are referred to laboratories in HICs, which is accessible to only a few patients in the South African public healthcare sector. Furthermore, referral abroad is not ideal for local patients, as genetic testing should ideally be tailored to the genetic architecture of local populations for maximum accuracy and efficacy. With the growing focus on non-communicable disease and genetic diversity on the African continent, it is also important to grow and develop genomic skills and capacity in Africa. It is thus imperative that the public health sector is able to offer a cost-effective NGS-based diagnostic protocol for paediatric epilepsy.

There are practical considerations to the implementation of such a laboratory testing protocol in the state health setting, which include but are not limited to cost, available expertise, facilities for processing, and computational resources. All these considerations are relevant when designing an NGS protocol for genetics-based diagnosis of paediatric epilepsy for translation into the South African state health sector if clinical translation to a resource-limited service environment is the ultimate goal[96].

1.9 Aims and Objectives

The main aim of this project was to develop an NGS-based sequencing protocol for the diagnosis of genetic paediatric epilepsies for translation into the testing repertoire of the National Health Laboratory Service (NHLS) in the Western Cape. The objectives were as follows:

a. To collect saliva samples (or blood when saliva is not possible) from children with refractory epilepsy and their parents at the epilepsy clinics at Red Cross Children's War Memorial Hospital and to extract DNA from those samples.

b. To choose an appropriate NGS approach which would be translatable to the South African State Healthcare Service, and design a testing method for genetically caused paediatric epilepsy based on that NGS approach.
c. To perform sequencing on the proband DNA using Ion Torrent Next Generation Sequencing technology.
d. To use a range of bioinformatics tools to establish or exclude the pathogenicity in the identified sequence variants.
e. To use cycle sequencing to confirm the presence of variants in the proband and to perform parental DNA analysis to determine the presence of *de novo* or inherited changes.
f. To assess the feasibility of the laboratory and bioinformatics process for potentially translation and implementation in the NHLS genetics laboratory.

The secondary aim was to gather preliminary pharmacogenomic data on a group of patients with refractory epilepsy as baseline data, to create a foundation for future research on the pharmacogenetics of refractory epilepsy in African populations, and to add to the limited pharmacogenomic data available for the South African population. The objectives were as follows:

a. To use establish which variants may influence ASM metabolism in study participants, making note of which ASMs are used in the South African State Healthcare Service and the pharmacogenomic database PharmGKB.
b. To design an appropriate approach to test for those variants in the study participants, or identify a commercially available test for those purposes.
c. To statistically compare the genotype distributions in the study group to genotype distributions general African populations.
d. To add the collected data to the pharmacogenomic data available for the South African population.

Chapter 2: Methods and Materials

2.1 Collection of samples and extraction of DNA

This research project was a sub-study of a wider parent study entitled “Precision management of epilepsy in South African children” (HREC 676/2017), which aimed to investigate how home technology-based monitoring may have the potential to improve the management of children with epilepsy in South Africa, in combination with molecular-based diagnoses and pharmacogenomic analyses.

The recruitment process of the parent study included the collection of saliva samples from participants and their parents. Ethical approval for this sub-study was granted by the Human Research Ethics Committee of the Faculty of Health Science, University of Cape Town (HREC 357/2019). All participants attended the specialised paediatric epilepsy clinic at the Red Cross War Memorial Children’s Hospital (RCWMCH) in Cape Town, Western Cape, South Africa. Recruitment pamphlets were made available to participants in three different languages (Afrikaans, English, and isiXhosa) (Appendix A). Sample collection was performed on participants and both their parents, where possible, after obtaining written informed consent.

All participants recruited into the parent project were included in this molecular study. Recruitment was based on the following criteria:

Inclusion criteria:

i. A phenotype of clinically defined complex epilepsy with pharmacoresistance, defined as failure to control seizures despite the prescription of two appropriate ASMs.
ii. Aged between 4 and 16 years at the time of recruitment. The reasons for this criterion were practical, as sampling involved collection of saliva and wearing of a watch to monitor physiological symptoms.

No participants were excluded on the basis of ethnicity or sex. A total of 40 participants were recruited, including ten child-parent trios, 23 child-parent duos (all child-mother) and seven participants with no recruited parents.

Saliva specimens were collected for DNA extraction, as saliva collection is minimally invasive and traumatic, especially for children. Samples were collected using the Oragene Saliva and Assisted Saliva kits (DNA Genotek, Ottawa, Canada). DNA was extracted according to the manufacturer's instructions for extraction from 4mL of saliva (Appendix B). DNA quantification and measurement of the 260/280 and 230/280 ratios, to assess the DNA purity and quality, were performed using the NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

2.2 Selection of the analytical approach

Next-generation sequencing is currently the most effective and widely used technology for sequence variant detection in heterogeneous disease such as epilepsy. When deciding between sequencing approaches, consideration was given to both the technical and practical aspects of ES and gene panels, as well as the clinical and ethical considerations of both approaches (Table 1). Consideration was given to the following factors:

i. Choice of the NGS chemistry and the capabilities of the NGS platform available.
ii. Affordability, such as reagent cost and cost of sequencing at a diagnostically appropriate depth of coverage (DOC).
iii. Requirements and availability of additional instrumentation and/or analytical software.
iv. Data storage and required computational capacity.
v. Options for sequencing data re-analysis at a later date.
vi. Availability of bioinformatic support.
vii. Ethical considerations, such as incidental findings and the capacity to address them, e.g., special consent, the need for an internal policy on handling incidental findings, etc.

Given the resource limitations of the South African state healthcare service, logistical and cost-related aspects were a major consideration in proposing a functioning and feasible NGS laboratory protocol. The approach ultimately decided on for this study was a gene panel on the Ion Torrent Personal Genome Machine (PGM) in the Division of Human Genetics at the University of Cape Town (UCT), as the currently most easily accessible instrument to the local diagnostic laboratory.

Table 1: Comparison of Exome Sequencing and Gene Panel diagnostic approaches

Characteristic of Sequencing Approach	Exome Sequencing	Gene Panel
Platform	Requires an NGS instrument capable of ES, which may affect scalability and cost. Panels still possible but may require higher throughput for cost-effectiveness.	Can be run on smaller NGS platforms, which allow better scalability and are more economical for low throughput environments. Not capable of human ES or GS.
Depth of Sequencing and Coverage	Lower on average than gene panels. May be challenging to reach clinical threshold for diagnostic purposes. Higher rate of exonic/allele drop-out than targeted gene panels.	Produces higher depth of coverage and on average, better gene coverage than ES. Regularly exceeds threshold required for diagnostic purposes. Markedly lower drop-out rates than ES.
Reanalysis of Data for New Candidate Genes	Allows for reanalysis of data in light of new gene discoveries.	Reanalysis of new genes is not possible. Data produced only for the genes included in the panel.
Cost	On average more expensive than gene panels, as produces more data.	More affordable than ES.
Data Storage and Analysis Requirements	Requires extensive data storage facilities. Requires sophisticated data analysis approaches.	Less data produced. Both analyses and data storage are more accessible and easily implemented.

Personnel and Expertise Requirements	Implementation of the bioinformatic pipelines requires highly trained personnel with particular expertise.	More accessible technique. Does not require as extensive training in order to perform process and analyses.
Incidental Findings	May result in incidental findings. Requires set up of laboratory and consenting protocols for reporting/not reporting.	No or minimal incidental findings.

2.3 Epilepsy gene panel design

The genes selected for inclusion in this panel were chosen via a sequential process, which involved several rounds of selection, each based on different criteria (Figure 1). Throughout the design process, the ultimate aim was to produce a panel which would:

a) include the genes commonly implicated in paediatric-onset epilepsies (especially the neonatal/infantile-onset DEEs), but

b) also be broad enough to accommodate a range of phenotypes,

c) include genes currently known to have precision medicine implications, and

d) be of manageable size for efficient, affordable, and easy implementation in the local diagnostic setting.

Initially, genes were identified through review of epilepsy gene panels available in HICs. The panels surveyed, included those offered by clinical genetic testing companies like GeneDX (Gaithersburg, Maryland, USA) and Invitae (San Francisco, California, USA), as well as a gene panel currently used in research investigating the genetic aetiology of the DEEs at UCT. This process resulted in a list of 164 genes for possible inclusion in the panel.

Genes identified from those panels were then reviewed using the existing literature for the associated phenotypes (e.g., childhood-onset or adult-onset epilepsy, metabolic or

neurological disorders, etc.), function (if currently described), level of association (i.e., proven association with multiple lines of evidence), as well as yield of informative findings and the availability of precision treatment for variants in those genes[24,97]. The third round of selection involved review of recent literature and online resources for genes, as well as the review of the selected genes using the ClinGen Expert Panel resource for epilepsy gene curation (https://clinicalgenome.org/affiliation/40005/)[24,98] . A summary of the gene selection process may be found in Figure 1. A final list of 78 genes (Table 2) was selected to design an Ion AmpliSeq™ Panel from Thermo Fisher Scientific (Foster City, CA, USA) (hereafter, "Thermo Fisher"), using the online interface for panel design (http://ampliseq.com) linked to the Thermo Fisher Scientific website (http://www.thermofisher.com/za).

The final panel design was based on the hg19 build of the human genome, with an amplicon range of 125-275 base pairs (Table 3). The 78-gene panel designed for this project consisted of 73 genes described by the manufacturer as "On Demand" and five genes incorporated into a separate, so-called "spike-in" panel. The "On Demand" genes are genes that have been optimised and pre-tested by the manufacturer to maximise the gene coverage and capture, and increase the success rate of the sequencing runs. The Ampliseq™ On Demand Panel may only be ordered for a relatively small number of reactions (with the number of reactions orderable being multiples of 96, 32, and 8) and so the necessity of reordering the panel allows for its redesign and/or incorporation of newly discovered/implicated genes.

The spike-in panel comprised of five genes not available "on-demand" (*ARX, FOXG1, GRIN2D, NR2F1,* and *PURA*) (Table 3). That is, they were not optimised and pre-tested and had to be ordered as a separate mini-panel ("spike-in panel"). In summary, an optimised, "On Demand" Ion AmpliSeq™ panel of 73 genes and a spike-in panel of 5 genes was ordered from Thermo Fisher.

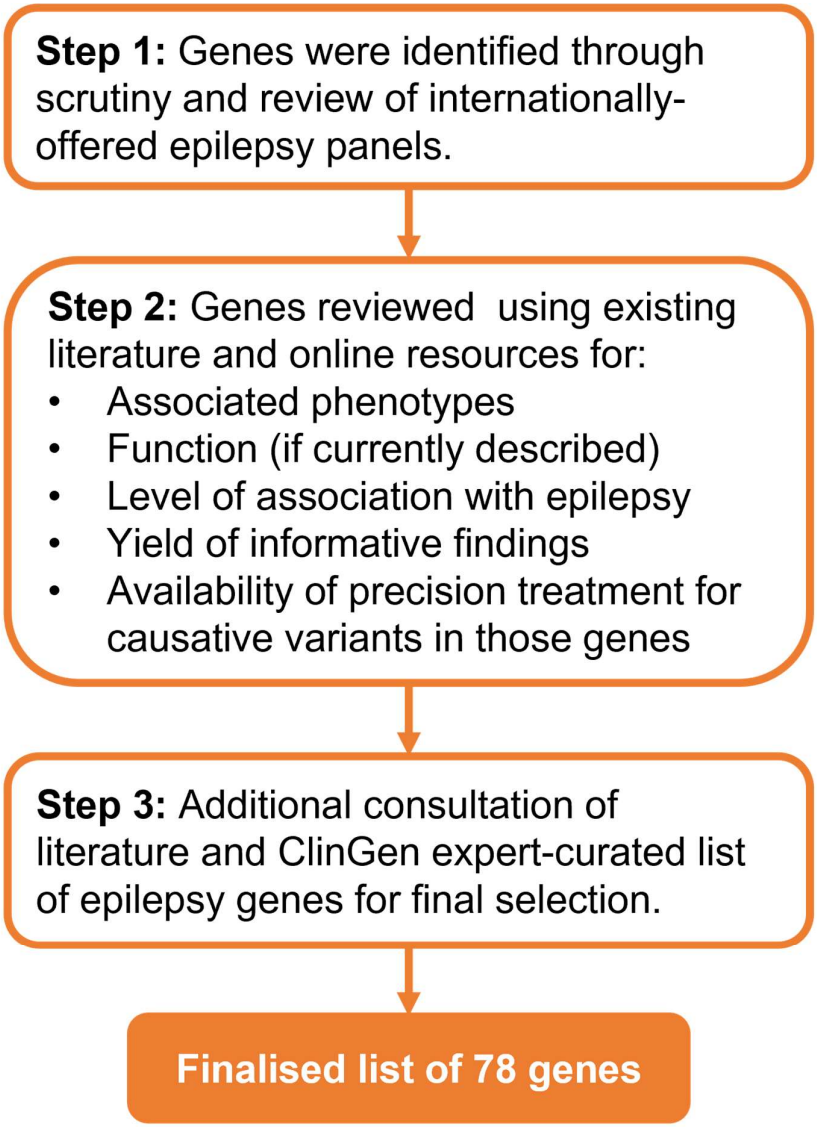


Figure 1 Summarised process of gene selection for panel design.

Table 2 Genes incorporated in gene panel design (OMIM reference number).

ALDH7A1 (107323)	*GABRB1* (137190)	*MECP2* (300005)	*SLC35A2* (314375)
ALG13 (300776)	*GABRB2* (600232)	*MEF2C* (600662)	*SLC6A1* (137165)

*ARX** (300382)	*GABRB3* (137192)	*NPRL2* (607072)	*SLC6A8* (300036)
ATP1A2 (182340)	*GABRG2* (137164)	*NPRL3* (600928)	*SMC1A* (300040)
ATP1A3 (182350)	*GNAO1* (139311)	*NR2F1** (132890)	*SNAP25* (600322)
CACNA1A (601011)	*GRIN1* (138249)	*PCDH19* (300460)	*STX1B* (601485)
CACNA1E (601013)	*GRIN2A* (138253)	*PNKP* (605610)	*STXBP1* (602926)
CASK (300172)	*GRIN2B* (138252)	*PNPO* (603287)	*SYN1* (313440)
CDKL5 (300203)	*GRIN2D** (602717)	*POLG* (174763)	*SYNGAP1* (603384)
CHD2 (602119)	*HCN1* (602780)	*PRRT2* (614386)	*SZT2* (615463)
CLCN4 (302910)	*HNRNPU* (602869)	*PURA** (600473)	*TBC1D24* (613577)
COL4A1 (120130)	*IQSEC2* (300522)	*SCN1A* (182389)	*TBL1XR1* (608628)
DCX (300121)	*KCNA2* (176262)	*SCN1B* (600235)	*TCF4* (602272)
DEPDC5 (614191)	*KCNB1* (600397)	*SCN2A* (182390)	*TSC1* (605284)
DNM1 (602377)	*KCNH5* (605716)	*SCN3A* (182391)	*TSC2* (191092)
DYRK1A (600855)	*KCNQ2* (602235)	*SCN8A* (600702)	*UBE3A* (601623)
EEF1A2 (602959)	*KCNQ3* (602232)	*SIK1* (605705)	*WDR45* (300526)

*FOXG1** (164874)	*KCNT1* (608167)	*SLC13A5* (608305)	*WWOX* (605131)
GABBR2 (607340)	*KIAA2022* (300524)	*SLC1A2* (600300)	
GABRA1 (137160)	*MBD5* (611472)	*SLC2A1* (138140)	

Note "*" denotes genes which are not available as "On Demand" genes and were ordered as a separate "spike-in" panel; underlined genes are genes which currently have specific precision medicine treatment implications[24,46].

Table 3: Details of panel design, taken from the panel design interface at http://ampliseq.com

	Ion AmpliSeq™ On Demand Panel	**Spike-In Panel**
Genome	Human (hg19)	Human (hg19)
Amplicon Range	125-275 bp	125-275 bp
Panel Size	389.953 kb	9.4kb
In-Silico Coverage	99%	71.89%
Number of Primer Pools	2	2
Pool 1	1018 amplicons	28 amplicons
Pool 2	1014 amplicons	25 amplicons

The coverage of the genes included in the spike-in panel was predicted by Thermo Fisher through an *in silico* analysis, and provided upon the completion of the panel design. Although the predicted coverage was less than 90% for all genes (Table 4), the spike-in panel remained included in the analysis due to the importance of including those genes, and the fact that the predicted coverage was based on an *in silico* analysis, and may have performed better in practice than predicted. (Also of note, only the coding DNA sequence (CDS) region of genes is covered by the On-Demand panel design tool, thus UTR regions are not specifically included. The exon padding for every gene design is either 5 bp or 25 bp on the 5' and 3' ends.)

Table 4: Predicted spike-in panel design gene coverage based on *in silico* analysis, provided by Thermo Fisher Scientific

Gene	Total Number of Targeted Base Pairs	Base Pairs Predicted to be Covered	Base Pairs Predicted to be Missed	Predicted Total Coverage
ARX	1739	1228	511	70,62%
FOXG1	1480	927	553	62,64%
GRIN2D	4131	2993	1138	72,45%
NR2F1	1302	905	397	69,51%
PURA	979	871	108	88,97%

2.4 DNA quality check assays

Stock DNA was diluted in nuclease free water to a concentration of 50ng/uL. The integrity of the diluted DNA was checked by agarose gel electrophoresis, using 1% agarose gels. Samples dilutions were loaded onto gels as 1uL of each sample dilution with 3uL of loading dye, and 8uL of Gene Ruler 100bp Plus DNA Ladder, followed by electrophoresis at 120V for 40 minutes. Samples which did not show on the gel were rerun with 2uL of DNA dilution (100ng of DNA) loaded onto the gel.

A real-time PCR RNase P amplification assay was performed as a quality control check to both quantify and assess the efficiency of DNA amplification for each sample. The assay was performed using the TaqMan® RNase P Detection Reagents Kit (Thermo Fisher Scientific, USA (catalogue number: 4316831)), according to manufacturer's instructions (Manual Title: *Demonstrated Protocol: Sample Quantification for Ion AmpliSeq™ Library Preparation Using the TaqMan® RNase P Detection Reagent Kit*; Publication Number: MAN0007732).

The construction of a standard curve using DNA of known concentrations is necessary in order to quantify sample DNA amplification during the PCR RNase P amplification assay. Prior to preparation and quantification of patient samples, solutions of known DNA concentration were

prepared with control DNA, provided with the RNase P Detection Reagents kit. Seven tubes were prepared, and labelled 1 to 7. 15uL of nuclease-free water was added to each tube. A sequential 1:2 dilution was performed: 15uL of control DNA added to tube 1 and mixed by vortexing, 15uL of solution from tube 1 was added to tube 2 and mixed by vortexing, with the process repeated until the dilution series of seven standard samples was completed, for the following concentrations in each tube:

Table 5: DNA concentrations used to construct a standard curve to quantify sample DNA amplification.

Tube	Volume (uL)	Concentration (ng/uL)
1	15	5
2	15	2.5
3	15	1.25
4	15	0.625
5	15	0.3125
6	15	0.15625
7	30	0.078125

Patient samples and control DNA solutions were diluted in 1:100 and 1:500 dilutions in nuclease-free water. A PCR master mix was prepared using TaqMan® Universal PCR Master Mix (2X), 20X RNase P Primer-Probe mix, and nuclease-free water. 17.5uL of master mix and 2.5uL of either patient sample dilution or control DNA dilution were added to wells in a 96 well plate, with each individual sample being added three times for readings in triplicate. Nuclease-free water was used as a non-template control. Thermocycling was carried out using a Bio-Rad C1000 Thermo Cycler and a CFX96 Real Time System (Bio-Rad Laboratories, Hercules, California, USA). Cycling conditions were as follows: 2 minutes at 50°C; 10 minutes at 95°C, then 40 cycles of 15 seconds at 95°C, followed by 1 minute at 60°C.

DNA showing poor amplification with the RNase P assay were subjected to DNA clean-up with the OneStep™ PCR Inhibitor Removal Kit or the DNA Clean and Concentrator™ Kit (both

manufactured by Zymo Research (Irvine, California, USA)), with four samples cleaned with each kit. Clean-up was not repeated with the other 32 samples due to substantial reduction of DNA concentration.

2.5 Next-Generation Sequencing

Next-generation sequencing was performed on 40 proband samples using the Ion Torrent™ technology PGM platform (Thermo Fisher Scientific) in the Division of Human Genetics, Faculty of Health Science, UCT. The Ion Torrent™ technology sequences DNA through "sequencing by synthesis" (also described as ion semi-conductor sequencing). The release of hydrogen ions during incorporation of new nucleotides results in changes in hydrogen ion concentration (that is, pH level). The difference in hydrogen bonding between guanine and cytosine and adenine and thymine base pairs results in changes to the hydrogen ion concentrations, which can be monitored and used to determine the sequence of the nucleotides being added.

2.5.1 Library preparation

In the context of the Ion Torrent™ NGS technology, library preparation refers to the multi-step process of preparing DNA samples for sequencing. Preparation involves the fragmenting of DNA, ligating "adaptors" to the DNA and enrichment for the targeted regions by amplification (when using on the Ion Torrent™ platform). The non-target DNA is removed. Adaptors are nucleotide sequences which enable emulsion PCR, which refers to the amplification of DNA in separate individual water-in-oil drops, to avoid the formation of PCR artifacts[99].

Each DNA sample was diluted to 1ng/uL, with volumes calculated according to concentrations measured by the real-time PCR assay described above. Automated library preparation was performed for each sample using the Ion Chef™ Instrument (Thermo Fisher Scientific) and the Ion AmpliSeq™ Kit for Chef DL8 kit (Thermo Fisher Scientific), according to manufacturer's instructions (Manual Title: *Ion AmpliSeq™ Library Preparation on the Ion Chef™ System*; Publication Number: MAN0013432). The Ion Chef™ Instrument is a machine which performs automated library preparation, once samples have been prepared using the Ion AmpliSeq™ Kit for Chef DL8 kit. 15uL of diluted DNA samples were added to 96 well IonCode™ PCR

Plates, provided in the Chef DL8 kit. Reagents were thawed at room temperature for 20 minutes prior to use. Reagents, primer pools used to enrich (through amplification) the regions targeted for sequencing, the DNA-containing plate, and a collection tube were loaded onto the machine. Samples underwent automatic library preparation in sets of eight, the maximum number of samples which can undergo library preparation at one time on the Ion Chef™ Instrument. Annealing and extension time was set to four minutes for all samples (as dictated by the number of primer pools in the gene panel, according to the manufacturer's instructions).

Prior to loading of samples on the Ion Chef™ Instrument, the spike-in primer pools were added to the parent panel (consisting of the On Demand genes) primer pools. The volume of spike-in primer pools to be added to the parent panel is dictated by the number of primer pairs in the primer pools in the spike-in panel; 6.4uL of each spike-in primer pool was added to its corresponding parent panel pool, as per the manufacturer's instructions. No additional quantification of the prepared library was required, as automatic library preparation produces libraries at a uniform concentration of 700 picomole (pM).

2.5.2 Templating of DNA onto Ion Sphere Particles

Upon the completion of library preparation, sample amplicons must be attached to microscopic beads to undergo clonal amplification prior to actual sequencing. This process, and the loading of the samples onto a so-called "chip" for sequencing, is referred to as "templating". It is necessary to determine in advance how many samples should be loaded onto one chip, for the most economical use of the reagents, while considering the minimum DOC required for the sequencing's purpose. It was determined that in order to achieve an average DOC of at least X300, eight samples should be loaded per one Ion 318™ Chip[100].

The concentration of sample libraries loaded onto a chip must also be determined prior to templating. Optimal library concentration for loading is influenced by the Ion Sphere particle (ISP) loading and levels of polyclonality, and finding the best balance between the two to produce "useable reads". ISP loading refers to the amount of DNA amplicons attached to an ISP bead, which influences the amount of data produced by the sequencing run. Too great an amount of ISP loading increases the likelihood of polyclonality occurring. Polyclonality refers to the attachment of more than one type of amplicon to an ISP bead, resulting in unusable

data, due to the production of mixed signals during the ion semiconductor sequencing. An appropriate balance between the two is thus required to produce the maximum amount of useful and reliable sequencing data. Determining that balance is a key part of panel optimisation, as there is no definitive threshold recommended by the manufacturer, rather each panel must be individually optimised based on its characteristics and purpose.

In order to determine the optimal conditions for the panel's use in the clinical sphere, individual chips (corresponding to five different sequencing runs) were loaded with different concentrations of sample libraries. Sample libraries were diluted from 700pM to a designated concentration of 30pM for one run, and 35pM for the other four runs using nuclease-free water. Templating of libraries on to Ion 318™ Chips was performed using the Ion Chef™ Instrument and Hi-Q™ Templating Kits, according to manufacturer's instructions (Manual Title: *Ion PGM™ Hi-Q™ View Chef Kits User Guide*; Publication Number: MAN0014571). Once templating was complete, chips were immediately removed from the Ion Chef™ Instrument and placed directly for sequencing on the Ion Torrent PGM, or placed in an airtight container and refrigerated until sequencing.

2.5.3 Sequencing

Sequencing of samples was performed using the Ion Torrent™ PGM™ System (Thermo Fisher Scientific) (Manual Title: *Ion PGM™ Hi-Q™ View Chef Kits User Guide*; Publication Number: MAN0014571). Templated chips were placed in the PGM, and the number of flows was set to 500 ("flow" refers to the process of adding nucleotides to the sequencing reaction to allow for strand extension, i.e., sequencing by synthesis.) Five sequencing runs of 8 samples each were performed, for a total of 40 samples sequenced.

2.5.4 Data analysis and variant prioritisation

Upon the completion of the sequencing process, each proband was subjected to a variant prioritisation workflow to prioritise variants for further investigation. Run progress, finalised run information, and quality control metrics were analysed using the Torrent Suite™ software package (Thermo Fisher Scientific) in order to confirm the quality and amount of useable sequencing data. Sequencing run data was uploaded from Torrent Suite™ to the online

analysis tool Ion Reporter™ software (version 5.12) (Thermo Fisher Scientific) (http://ionreporter.thermofisher.com). The alignment algorithm used by the Torrent Suite™ software to align reads to the reference genome (hg19) was TMAP, a currently unpublished algorithm[101,102].

A custom workflow was designed for analysis of the sequencing data. The workflow was designed (through the selection of various options available on the Thermo Fisher Ion Reporter™ Software) to:

i. Detect and annotate variants in human DNA samples.
ii. Analyse a single sample at a time.
iii. Compare the sample sequences to the hg19 reference genome (as these sequencing runs were performed at the end of 2019 and throughout 2020).
iv. Focus on the target regions specific to the panel (using the .bed file generated by Thermo Fisher Scientific upon submission of the panel design to specify those regions for the software).
v. Perform annotations using all available annotation sets:
    - Locus, variant type, population frequencies, gene, transcript, genomic location, function, coding change, *in silico* predictions of pathogenicity, designation in ClinVar, designation in dbSNP, genotype ontology terms, and OMIM number.
vi. Have no automatic variant filters.
vii. Not call CNVs. (This requires baseline data, which was not available at this stage.)
viii. Not use any additional plug-ins (i.e., additional analyses, such as methylation analysis or automated variant calling).

Upon the completion of analysis for each sample, a .tsv file was exported from Ion Torrent for the relevant sample. These .tsv files were generated by the software using a .bam file (*Binary Alignment Map file* - the file containing the aligned sequence data), a .bai file (the index file which must necessarily accompany a .bam file), and a .vcf file (Variant Call Format file, containing all the sequence variations from the reference genome). After being exported, the .tsv files were transferred to Microsoft Excel (Microsoft Corporation, Redmond, Washington, USA), producing a spreadsheet of variants annotated with the parameters chosen as part of the Ion Reporter™ custom analysis workflow design (above).

The following filters were applied to the spreadsheet to prioritise variants for further investigation:

i. Prior to detailed filtering, a broad check was performed for any recurrent variants already published as pathogenic or likely pathogenic in ClinVar. These variants would be noted, but the extended variant prioritisation process would be resumed.

The reason for this initial check, was to ensure that no previously published pathogenic or likely pathogenic variants were filtered out because of position or function (e.g., intronic or synonymous variants with published evidence of pathogenicity).

ii. The initial filter applied to all the variants was the automatic assessment of the quality of reads produced in the sequencing data – that is, the raw data identifying the presence of a variant had to pass the Ion Torrent™ quality filters (marked as PASS), which is bioinformatically assessed by the software.

iii. All variants were then filtered based on their population frequencies in the 5000 Exomes project. The frequency cut-offs were 1) absence from population databases for AD genes, and 2) <5% minor allele frequency for variants in AR genes.

Exclusion of AD variants with published MAFs in population databases was based on the assumption that variants causative of these early-onset, severe phenotypes will not feature in population databases of normal genomic variation. The majority of causative variants for these disorders occur *de novo* in the absence of family history or an affected parent. The <5% cut-off for AR genes was chosen based on the accepted cut-off for the MAF of rare, Mendelian variants[103]. Variants with population frequencies greater than those cut offs were discarded. X-linked genes were filtered using the same 5% minor allele frequency cut-off as autosomal recessive genes, though consideration was given to the specific gene and sex of the patient under investigation, as not all X-linked epilepsy genes follow typical X-linked dominant or recessive inheritance patterns.

iv. All intronic, and synonymous variants, as well as those in the UTR regions and regions upstream or downstream of the gene were excluded. Exceptions included

synonymous variants affecting splicing (the first/last codon of an exon) and known pathogenic variants previously described in the literature and clinical databases, such as ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), LOVD (https://www.lovd.nl/), or HGMD (https://www.hgmd.cf.ac.uk/).

Subsequent to the prioritisation process described above, quality indicator metrics were also assessed. Depth of coverage was assessed, and variants were discarded if the depth of coverage was below the diagnostic minimum of 30[104]. However, in cases where the depth of coverage was low, but the variants in question had been previously reported and had other good quality indicators, such as the proportion of reads containing the variant, the .bam data file would be used to visualise the sequence context of the amplicon using the Integrative Genomics Viewer (IGV) (https://software.broadinstitute.org/software/igv). Based on this, a discretionary decision would be made on whether or not the variant should be included for cycle sequencing confirmation. (In greater depth, allele balance refers to the percentage of reads in which a variant is observed in sequencing data relative to the wild type allele. For example, for a heterozygous germline variant in an autosomal dominant gene, the allele balance should be close to 50% (or a ratio of 0.5). Allele imbalance may indicate a technical artifact from the sequencing process, or may indicate possible mosaicism, which may be relevant in certain phenotypes.)

v. Finally, consideration was also given to the strength of association between the gene in which a variant was found and the phenotypic presentation of the proband under investigation, as multiple candidate variants may be detected per patient.

For example, in a patient clinically diagnosed with DS, a truncating variant in *SCN1A* may be initially prioritised over variants in other genes. This does not mean that the other variants would be discarded, rather, they would be investigated if the *SCN1A* variant were excluded (e.g., if it were present in an unaffected parent). Upon each proband undergoing the variant prioritisation process, those variants which passed all the above-described filters were selected for direct cycle sequencing confirmation. The variant prioritisation pathway is summarised in Figure 2 below.

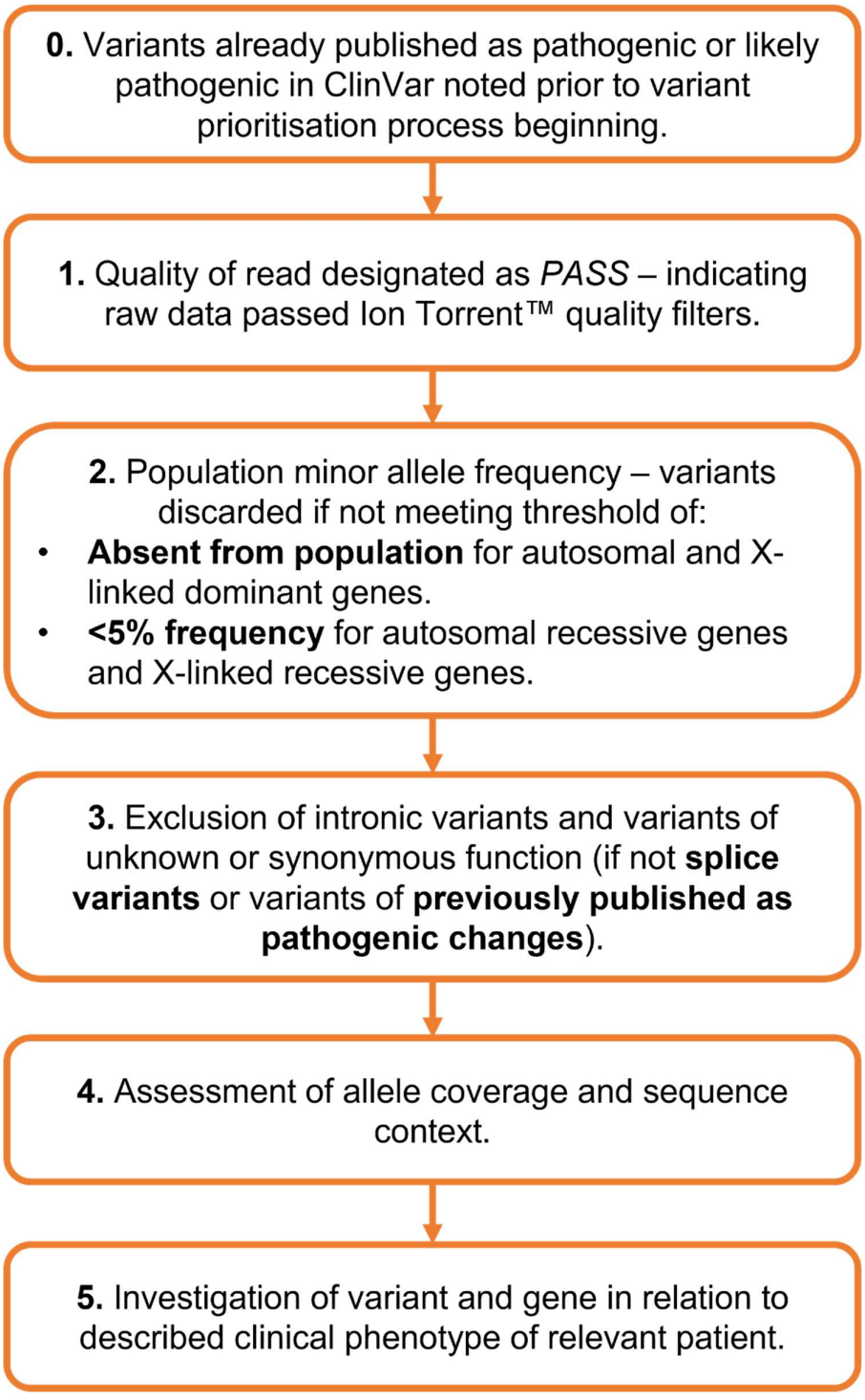


Figure 2 Summary of filtering pipeline to identify variants for further investigation.

2.6 Assigning pathogenicity classifications using ACMG criteria

Pathogenicity classifications were assigned to variants through use of the American College of Medical Genetics and Genomics (ACMG) guidelines for the interpretation of sequence variants[105]. The ACMG guidelines were developed to enable the use of standardised terminology and classification criteria for genetic variants, allowing for greater uniformity in how variants are classified and interpreted. Variants identified after sequencing and the application of the filtering workflow described above were systematically assessed across a range of criteria, such as *de novo* status, *in silico* predictions, alignment with expected phenotypes, population frequency, etc., to come to final classification. Such assessments required access to phenotypic information, which was accessed through the clinical REDCap database for the study.

2.7 Variant confirmation with direct cycle sequencing and segregation analysis

Following prioritisation, variants were selected for confirmation with direct cycle sequencing and segregation analysis. Pre-designed primers for direct cycle sequencing, specific to the regions of the particular genes in which the variants of interest are located, were ordered from Thermo Fisher Scientific. Sequencing was performed using the BigDye® Direct Cycle Sequencing Kit, according to the manufacturer's instructions (Protocol Title: *BigDye® Direct Cycle Sequencing Kit Protocol* (PN 4458040C) (Appendix C).

Parental DNA (when available) was tested for the presence of the same variant. Segregation analysis is an important step in variant interpretation, as establishing *de novo* occurrence is often a deciding factor determining variant pathogenicity in cases of DEE and other severe paediatric epilepsies. In these cases, *de novo* occurrence, together with clinical evidence, would override all *in-silico* predictions of pathogenicity. Final variant classification and relevance to phenotype and treatment were discussed with members of the clinical team responsible for the care of the patients.

2.8 Pharmacogenetics

The secondary aim of this project was to conduct a preliminary pharmacogenomic exploration in the study's participant group, which reflected the population of the Western Cape, South Africa.

2.8.1 Selection of single nucleotide variants for inclusion on a pharmacogenomic array

Within the context of minimal published research into the pharmacogenetics of ASMs, particularly in the African populations, the approach taken to initiate this preliminary exploration incorporated several lines of inquiry. The pharmacogenomic component of this study was not looking for disease causing variants, but rather was investigating potential associations between variants and drug metabolism, and hence consisted of a mainly statistical approach.

A comprehensive search for variants which may influence the absorption, distribution, metabolism, excretion, and transport of ASMs in African populations is well beyond the scope of this project. It is a research question which will be addressed by future research into the pharmacogenetics of ASMs in African populations.

It was within the scope of this project to consider the following:

1. What pharmacogenomic variants known to be implicated in the metabolism of other drugs may also be associated with absorption, distribution, metabolism, excretion, and transport (ADMET) of ASMs in this group of probands?
2. What variants considered to be specifically relevant to ASM metabolism in other populations may be present in this study group – that is, would previously identified variants associated with ASM metabolism be present in this research group?

Genotyping was performed with the use of pharmacogenomic arrays. Commercially available pharmacogenomic arrays were reviewed, and the VeriDose® Core Panel, designed by Agena

Bioscience (San Diego, California, USA), and a custom mass array (run on the MassARRAY system in combination with MALDI-TOF analysis) manufactured by Inqaba Biotec (Pretoria, South Africa) were selected.

The VeriDose® Core Panel consists of 73 genomic variants across 20 well-known absorption, distribution, metabolism, excretion (ADME) genes, including 68 single nucleotide variants (SNVs) or insertions/deletions (INDELs) and CNVs (See Appendix D for complete list of variants on the VeriDose® Core Panel). The VeriDose® Core Panel was thus utilised to investigate known ADME variants in the novel context of ASMs and an African population.

To investigate the presence of known ASM ADME variants in an African population, a custom mass array was designed to be run by Inqaba Biotec as the technical service provider. The online database PharmGKB (https://www.pharmgkb.org/) was used to identify variants in genes known to influence ASM metabolism in non-African populations. A sequential filtering process was followed to select the variants for inclusion. This process involved investigation of which variants influence the metabolism of ASMs commonly prescribed in the public health sector in South Africa (listed in Table 6), and review of the level of evidence for clinical annotation for each variant. (Clinical annotation here refers to descriptions in literature of associations between the variants and influence on the metabolism of each of the respective ASMs.) Levels of evidence are listed on PharmGKB as high, moderate, low, or preliminary evidence – only variants with high or moderate levels of evidence were considered for inclusion. The initial list of variants for inclusion in the custom mass array was compared to the VeriDose® Core Panel, and variants appearing on both lists were removed from the list for the custom mass array and only included in the VeriDose® Core Panel. The custom mass array was ultimately designed to investigate the presence of eight ASM ADME variants in the study population. Those eight variants are listed in Table 7 below.

Table 6 Anti-seizure medications commonly used in the South African Public Health Sector, administered to the participants in this study.

Carbamazepine	Lorazepam
Phenobarbital	Oxcarbazepine
Topiramate	Prednisone

Levetiracetam	Stiripentol
Sodium valproate	Vigabatrin
Valproic acid	Prednisolone
Lamotrigine	Methylphenidate
Clobazam	Risperidone
Phenytoin	Lorazepam

Table 7 Variants, and the genes in which they are found, included in this study

Gene	Variant
EPHX1	rs2234922
DRD2	rs1800497
DRD2	rs1799978
HTR2C	rs1414334
UGT1A4	rs2011425
EPHX1	rs1051740
SCN1A	rs3812718
MC4R	rs489693

2.8.2 Preparation of DNA samples for pharmacogenomic arrays

DNA from each proband was diluted based on the concentration measured with the NanoDrop Spectrophotometer to a concentration of 50ng/uL in a volume of 30uL. The prepared samples were subsequently referred to Inqaba Biotech for analysis using both the VeriDose® Core Panel and the custom array.

2.8.3 Analysis of pharmacogenomic data

Data from the VeriDose® Core Panel and the custom pharmacogenomic array were provided in the format of genotype counts in a Microsoft Excel (Microsoft, Redmond, Washington State, USA) spreadsheet. All subsequent analyses were carried out independently of Inqaba Biotec. Hardy-Weinberg Equilibrium (HWE) analyses were performed for each variant independently to assess if genotype and allele frequencies deviated from HWE. Deviation from HWE could suggest genotyping errors or a form of natural selection impacting genotype or allele frequencies. The Pearson Chi-squared test was performed for each variant in R (https://www.r-project.org/), using data for African reference populations available in the online database Ensembl (https://www.ensembl.org/), in order to investigate if the distribution of genotypes in the study population differed from that in an African reference population (the implication being if genotype distributions were significantly different that those variants may be involved in the abnormal metabolism of ASMs in the participant group). The Chi-squared test could not be performed for certain variants due to genotype distributions consisting of only one genotype present in a population, or due to a lack of data for a reference population.

Chapter 3: Results

3.1 Collection of samples and DNA extraction from saliva and blood

Forty participants matching the inclusion criteria were recruited from RCWMCH. Recruited participants displayed a wide range of symptoms and disease presentations, with a variety of working diagnoses suggested by clinicians (Appendix E). The study group comprised of ten complete patient-parent trios, 23 patient-parent duos (which were exclusively patient-mother duos), and seven patients for which no parental DNA was available. The study group comprised of 20 males and 20 females of the self-reported ancestries: Indigenous Black African (n = 28), South African Mixed Ancestry (n = 10), Indian (n = 1), and South African White (n = 1) (Table 1).

The concentrations of DNA extracted from the patients and parents ranged from 3.19ng/uL to 2644.62ng/uL (see patients in Table 1, both patients and parents in Appendix F). In cases where multiple samples were obtained from the same patient, the samples with the best quality indicators were processed moving forward into NGS. All patient DNA was successfully processed through NGS and pharmacogenomic analysis.

Table 1 Study participants, designated by study-assigned family numbers*, self-assigned ancestry, availability of parental samples, sample DNA concentrations and the quality indicator ratios.

Study-Assigned Number*	Self-Reported Ethnicity	Maternal Sample	Paternal Sample	DNA Concentration (ng/uL)	260/280 ratio	260/230 ratio
1	Mixed Ancestry	X		116.46	1.81	1.17
2	Black African			714.18	1.89	1.61
3	Black African	X	X	370.80	1.80	1.16
4	Mixed Ancestry	X		46.27	1.98	0.95

5	Black African	X		170.97	1.87	1.06
6	Black African			1648.45	1.90	1.50
7	Black African	X	X	329.73	1.84	1.29
8	Black African			172.73	1.91	0.79
9	Black African	X		240.51	1.68	0.82
11	Black African	X		1194.25	1.90	1.48
12	Black African	X		255.34	1.92	1.13
13	Black African	X	X	439.61	1.69	0.79
14	Black African			284.07	1.89	1.34
16	Black African	X		585.96	1.93	1.46
18	Black African			297.29	1.83	1.34
19	Mixed Ancestry	X		272.54	1.79	0.82
20	Mixed Ancestry			157.07	1.52	0.35
21	Black African	X	X	260.59	1.86	0.89
22	SA White	X	X	719.71	1.82	1.77
24	Indian	X	X	175.19	1.93	0.87
25	Mixed Ancestry	X		1065.33	1.85	1.57
26	Black African	X		481.00	1.70	0.98
27	Black African	X		419.40	1.94	1.38

28	Black African	X		2644.62	1.98	1.57
29	Black African	X		88.48	1.26	0.32
30	Black African	X	X	22.4	1.87	-2.78
31	Black African	X		279.38	1.65	0.65
32	Black African	X		484.21	1.87	1.37
33	Mixed Ancestry	X		90.70	1.74	0.87
34	Mixed Ancestry	X		341.01	1.83	0.75
35	Mixed Ancestry	X	X	277.18	1.79	0.98
36	Black African	X		221.37	1.95	0.91
41	Black African			128.38	1.60	0.60
42	Black African	X	X	1988.75	1.87	1.87
44	Black African	X	X	488.03	1.78	1.62
45	Black African	X		326.35	1.88	1.00
46	Mixed Ancestry	X		291.74	1.63	0.62
47	Black African	X		321.74	1.85	1.25
48	Mixed Ancestry	X		1156.83	1.87	1.27
50	Black African	X		491.72	1.86	1.06

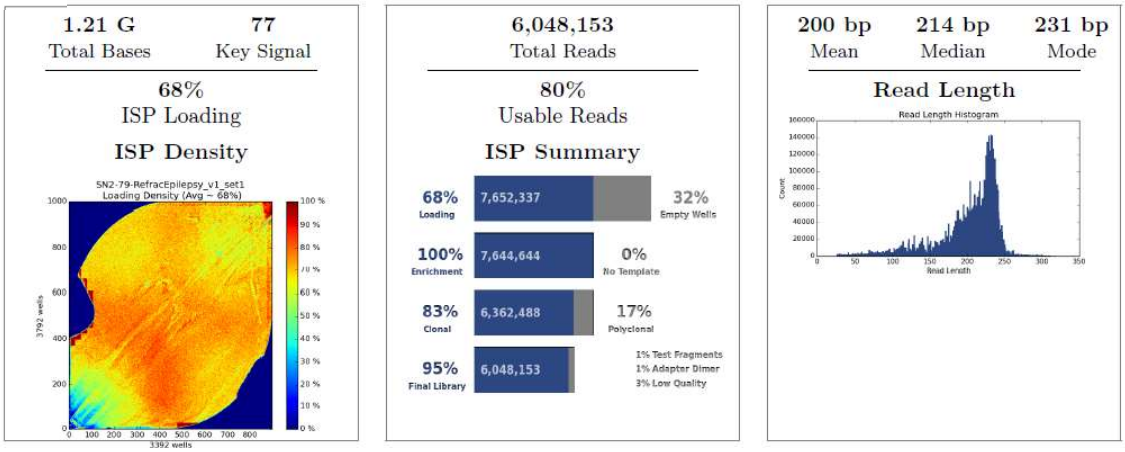
*Numbers were assigned sequentially, in order of recruitment. Missing numbers reflect families that withdrew or were removed from the parent study.

3.2 Gene panel

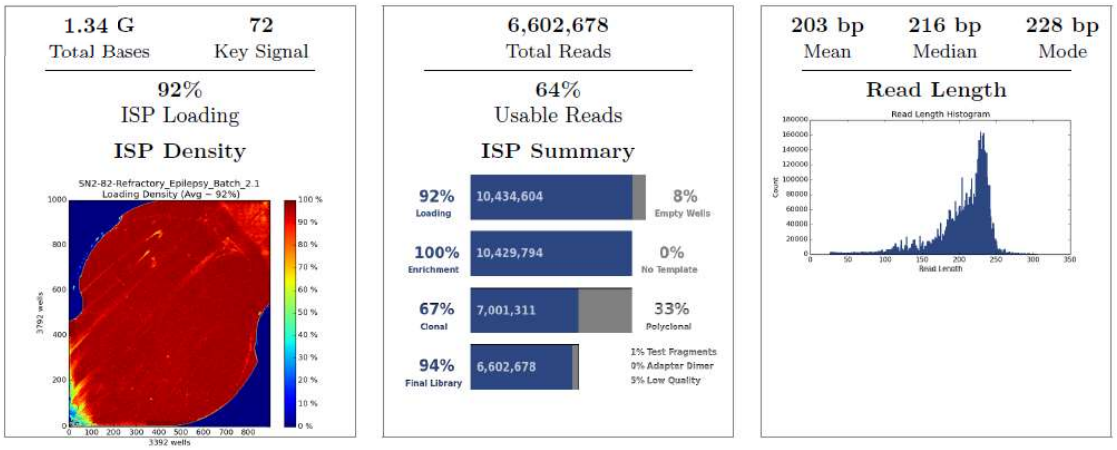
3.2.1 Gene panel technical performance and quality indicators

Based on a desired DOC for diagnostic purposes, ideal template loading was calculated to be eight samples run on an Ion 318™ chip, resulting in five sequencing runs of eight samples each, for a total of 40 probands sequenced. Templating of the initial run was done using 30pM of the prepared library, resulting in 80% usable reads (the percentage of which is determined by both Ion Sphere™ Particle loading and polyclonality as described in Chapter 2). The subsequent four runs templated 35pM of library onto the sequencing chip, achieving 64%, 69%, 54%, and 60% of usable reads (for runs 2 to 5 respectively). Figure 1A,1B, 1C, 1D, and 1E present the respective ISP density heat maps, summaries of usable reads, and summaries of read length for the five NGS runs performed, generated by the Ion Torrent software upon the completion of each run. All runs were completed successfully. Complete run reports for all runs, exported from Torrent Suite™ software, are available in Appendix G.

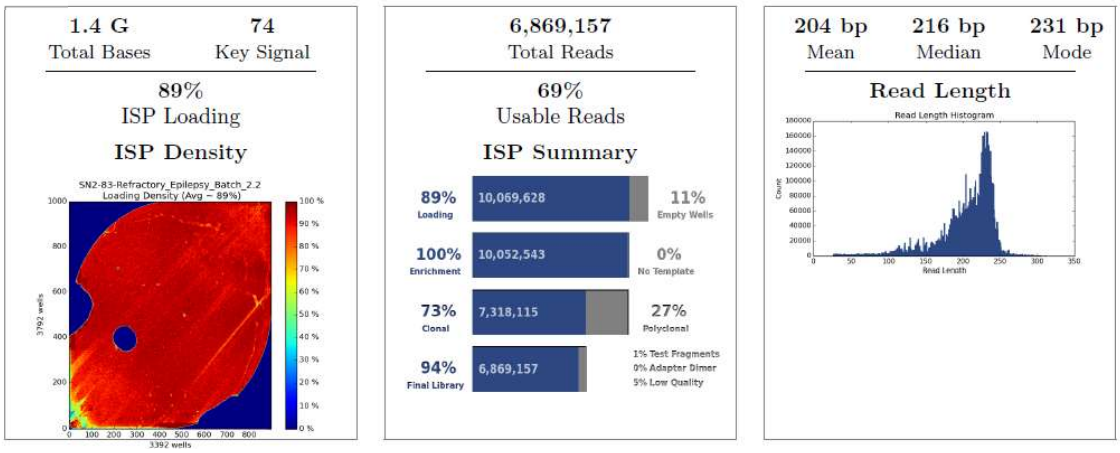
1A



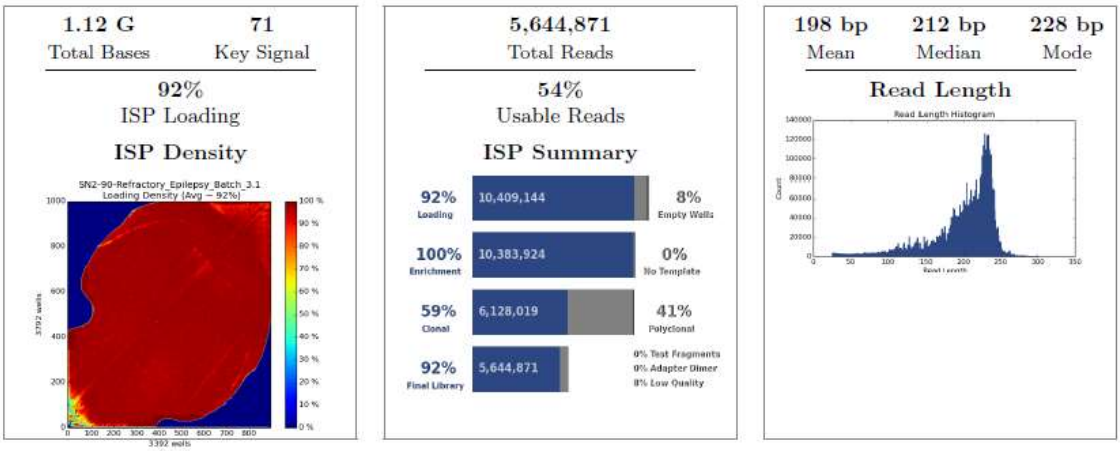
1B



1C



1D



1E

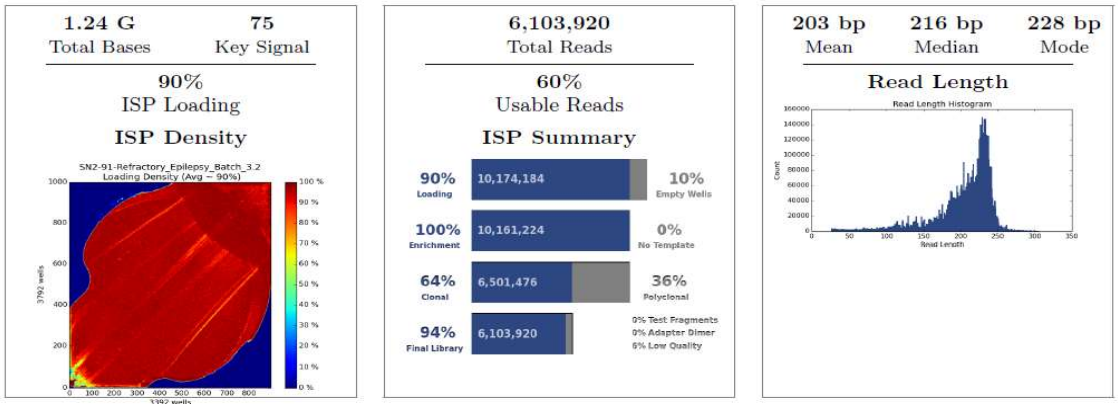


Figure 1 Summary of run data for Run 1-5 of the Refractory Epilepsy gene panel, produced by Ion Torrent software upon the completion of sequencing. Figure **1A** represents Run 1, Figure **1B** represents Run 2, Figure **1C** represents Run 3, Figure **1D** represents Run 4, Figure **1E** represents Run 5. (Full run reports available in Appendix G.)

3.2.2 Target Capture and Coverage

The average mean DOC across all targets in 40 probands was X363.51, ranging from X180.5 to X480.5. The mean sequencing depth across all target regions for each sample can be found in Table 2 below.

Table 2 Mean sequencing depth across all targets for each sample in sequencing runs 1-5.

Run 1	369.5	**Run 2**	368.3
	392.9		391.8
	407.7		425.5
	391.8		408.4
	392.1		384.2
	330.9		374.8
	206.2		373.1
	290.5		384.8
Run 3	418.7	**Run 4**	327.6
	462.8		322.9
	428.6		291.1
	315.8		348.1
	480.5		323.1
	370.3		297.4
	348.6		325.3
	420.5		335.6
Run 5	449.0		
	463.0		
	486.9		
	346.1		
	180.5		
	356.3		
	224.3		
	324.7		

Missing sequences were consistently detected by Torrent Suite™ software in four of the five genes in the spike-in panel, localised to exon 1 in *FOXG1,* exon 2 in *ARX*, exon 1 in *PURA*, and exons 2, 3, 4 and 13 in *GRIN2D* (Table 3). The regions surrounding the missing sequences were visually inspected in IGV in order to assess the genomic context which may have contributed to failed amplification (and thus, failed sequencing). Example images of the presentation of missing regions of sequence in IGV may be observed in Figure 2 below.

Table 3 Missing regions of sequence (drop-out) in the genes *FOXG1, ARX, PURA,* and *GRIN2D.*

Gene (NCBI RefSeq)	Site of Missing Sequence (hg19 Genomic Coordinates)	Genomic Context
FOXG1 (NM_005249.5)	Exon 1 (chr14: 29236486 – 29237041)	GC rich region – extended regions of G homopolymers or G/C dinucleotide repeats
ARX (NM_139058.3)	Exon 2 (chrX: 25031914-25031381)	GC rich region – stretches of G/C homopolymers
PURA (NM_005859.5)	Exon 1 (chr5: 139493767-139493996, and chr5: 139494220-139494463)	GC rich region – stretches of G/C homopolymers
GRIN2D (NM_000836.4)	Exon 2 (chr19: 48897914-48898193)	GC rich region – stretches of G/C homopolymers.
GRIN2D (NM_000836.4)	Exon 3 (chr19: 48901649-48901999)	GC rich region – stretches of G dinucleotide and trinucleotide repeats
GRIN2D (NM_000836.4)	Exon 4 (chr19: 48908133-48908265)	GC rich region
GRIN2D (NM_000836.4)	Exon 13 (chr19: 48945857-48946505, and chr19: 48946650-48947076)	GC rich region – stretches of G/C homopolymers.

Note The genomic context surrounding the missing sequence regions was observed by visually inspection of .bam files in IGV.

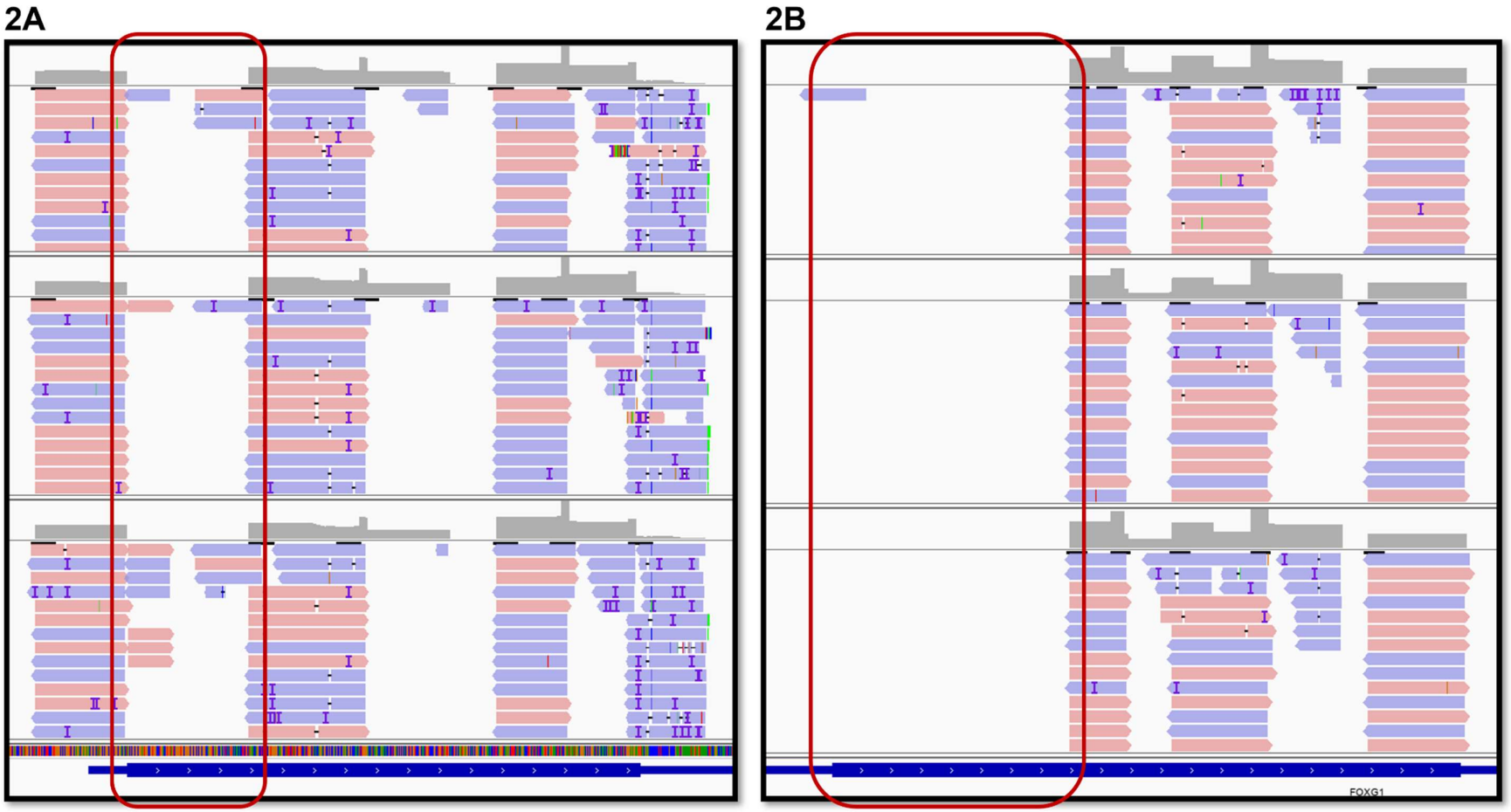


Figure 2 Non-exhaustive examples of missing regions of sequence in **2A** *PURA* and **2B** *FOXG1,* resulting in partial coverage of the respective exons. Sequence data from different patients is separated by double grey lines, region of missing sequence is highlighted by red boxes. Note the manner in which, across different patients, there is presence of some sequence to varying degrees before a total absence of reads.

3.2.3 Variant prioritisation and confirmation results

Once sequencing was complete, the .tsv files of all probands were combined and subjected to batch variant filtration as described in Chapter 2. After filtering, 17 variants were identified as variants of interest that warranted further investigation. Of these, 12 were confirmed by direct cycle sequencing and five were not observed on cycle sequencing, thus deemed to be sequencing artifacts (Table 4). All 12 confirmed variants were identified in dominantly inherited genes. Seven variants were observed in samples from non-affected parents during segregation analysis, and thus discarded as benign (Table 4). The five variants absent from (the available) parental samples were classified as pathogenic (n = 3, two in *SCN1A*, one occurring in a patient-parent trio, and the other in a patient-parent duo, and one in *GRIN2A*, occurring in a patient-parent duo), and as VUSs (n = 2, one in *GABRG2,* and one in *GRIN2B,* both occurring in patient-parent duos) (Table 5). The *in silico* indicators of pathogenicity for those five variants provided by the Ion Reporter software may be found in Appendix H. The filtering out of variants after each analytical step is summarised in Figure 3. Observed genotypes were aligned to the phenotypes of the patients through review by a multidisciplinary team. Detailed phenotypic descriptions of the patients in which those variants were found may be found in Table 6.

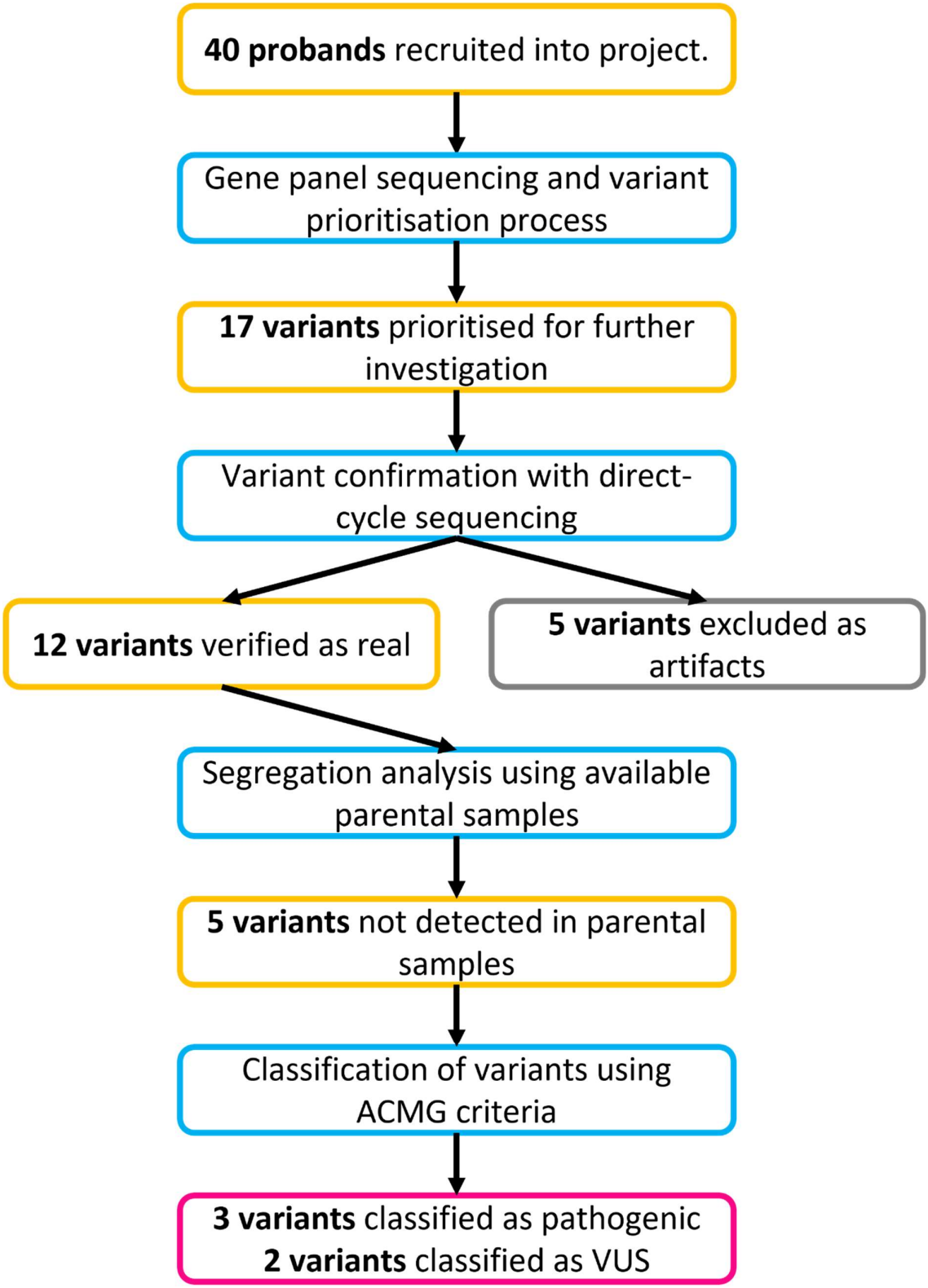


Figure 3 A flow diagram of the drop-off of variants through NGS variant filtration and prioritisation, direct cycle sequencing, and segregation analysis.

Table 4 Variants discarded as artifactual or classified as likely benign or benign

#	Gene (NCBI RefSeq)	Variant (hg19 Genomic Coordinate)	Type of Variant	Detected on Cycle Sequencing	Segregation Analysis
1	*UBE3A* (NM_130839.5)	c.959A>G (chr15: 25616362)	Missense	No - artifact	N/A
2	*IQSEC2* (NM_001111125.2)	c.1629delC (chrX: 53280128)	Frameshift Deletion	No - artifact	N/A
3	*SYNGAP* (NM_006772.2)	c.3167delG (chr6: 33411491)	Frameshift Deletion	No - artifact	N/A
4	*GABBR2* (NM_005458.7)	c.1489_1490insT (chr9: 101151175)	Frameshift Insertion	No - artifact	N/A
5	*HNRNPU* (NM_031844.2)	c.2069_2070insA (chr1: 245019303)	Frameshift Insertion	No - artifact	N/A
6	*COL4A1* (NM_001845.5)	c.2059C>T (chr13:110835376)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.
7	*CASK* (NM_003688.3)	c.1186C>T (chrX: 41448815)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.
8	*GABRA1* (NM_001127645.1)	c.8A>G (chr5: 161277824)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.
9	*GRIN1* (NM_001185090.1)	c.1394C>T (chr9: 140055641)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.

10	*DEPDC5* (NM_001242896.2)	c.2143G>T (chr22: 32229939)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.
11	*ALG13* (NM_001099922.2)	c.1955G>A (chrX: 110970262)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.
12	*ALG13* (NM_001099922.2)	c.1955G>A (chrX: 110970262)	Missense	Yes	Dominantly inherited gene, and detected in unaffected parent.

Table 5 Details, applicable ACMG criteria, and ACMG classification of variants

Patient Identifier	Gene (NCBI RefSeq)	rs ID and ClinVar ID	Variant Type	Coding Change (hg19 Genomic Coordinate)	Amino Acid Change	Zygosity	Inheritance	Applicable ACMG Criteria[122]	ACMG Classification[122]
34	*SCN1A* (NM_006920.5)	rs1553521567 ClinVar ID: 530456	splice acceptor	c.4444-1C>T (chr2:166852628)	N/A	Het.	Not established	PVS1 – Splice site variant PM2 – low population frequency (ref) PP4 – expected phenotype PP5 – described as pathogenic in online databases	Pathogenic
22	*SCN1A* (NM_006920.5)	rs121918624 ClinVar ID: 12889	nonsense	c.664C>T (chr2:166909392)	p.Arg222Ter	Het.	*de novo*	PVS1 – nonsense mutation PS2 – *de novo* mutation PM2 – low population frequency PP4 – expected phenotype	Pathogenic

								PP5 – described as pathogenic in online databases	
31	*GRIN2A* (NM_001134407.2)	rs796052549 ClinVar ID: 205657	missense	c.2191G>A (chr16:9892299)	p.Asp731Asn	Het.	Not established	PS3 – Published functional studies supportive of damaging effect on gene or gene product PM2 – low population frequency PM5 – A different missense variant at the same site is pathogenic PP2 – missense variant PP3 – computational evidence for pathogenicity	Pathogenic
16	*GABRG2* (NM_198903.2)	n/a	missense	c.211A>G (chr5:161520937)	p.Asn71Asp	Het.	Not established	PM2 – low population frequency PP2 – missense variant	VUS

5	*GRIN2B* (NM_000834.4)	rs1042339 ClinVar ID: 245684	missense	c.3499G>A (chr12:13716673)	p.Val1167Ile	Het.	Not established	PM2 – low population frequency PP2 – missense variant PP4 – expected phenotype	VUS

NOTE “Het.” is an abbreviation for heterozygous. No rs ID or ClinVar ID are available for *GABRG2* c.211A>G as it has not been previously described or submitted to online databases.

Table 6 Detailed phenotypic description of participants in which relevant variants (Table 5) were found

Detected Variant (hg19 Genomic Coordinate)	Age-at-onset	Seizure description	ASMs and response[1]	Development prior to seizure onset	Current development	Other clinical features/ conditions	EEG	Imaging	Working Diagnosis prior to Genetic Testing
SCN1A c.4444-1C>T (chr2:166852628)	2 months	Initially generalised tonic clonic, followed by myoclonic jerks. Right focal seizures at 2 years of age with secondary generalisation. Atonic seizures and drop attacks listed, but not age of onset. 5-9 seizures a day	*Trialled ASMs:* Carbamazepine Clobazam Lamotrigine Levetiracetam Phenobarbital Sodium Valproate Topiramate Carbamazepine worsened myoclonus. Lack of seizure control so different combinations tried.	Developmental delay prior to seizures - not smiling by seizure onset at 10 weeks.	Currently has severe ID - no evidence of neuroregression.	Initially right hemiplegia, now spasticity. Microcephaly, prominent ears	Abnormalities described as bilateral independent spiking, more marked over left hemisphere with isolated right hemispheric slowing. EEG evolution described as bilateral slowing	CT and MRI - no observed abnormalities on scans	Early Onset Epileptic Encephalopathy
SCN1A c.664C>T (chr2:166909392)	5 months	Initially right focal seizures with secondary generalisation.	*Trialled ASMs:* Carbamazepine Ethosuximide Levetiracetam	Extent of developmental delay not clear.	Mild to moderate learning disability at LSEN school.	None noted	Occasional spindle formations and	MRI - no abnormalities	Dravet Syndrome

		Evolved to multiple seizure types: prolonged focal, unaware, GTC, myoclonic jerks. Seizures a few times a week.	Phenobarbital Sodium Valproate Topiramate Levetiracetam had no observable effect. Sodium valproate observed to exacerbate neuroregression		Evidence for neuroregression: early development normal - regression with clusters of seizures Loses some skills after a lot of prolonged seizures but recovers.		V waves in sleep. Rare discharges from both hemispheres with wide field spread but frontal predominance.		
GRIN2A c.2191G>A (chr16:9892299)	72 months	No real change in semiology since onset. Left sided clonic jerks followed by post-ictal sleeping. Seizures about once per day.	*Trialled ASMs:* Carbamazepine Clobazam Lamotrigine Levetiracetam Sodium Valproate Response not recorded.	Apparently normal development until onset of seizures, but extent of developmental delay is unclear from records.	No evidence of neuroregression.	Left hemiplegia. Possible cerebral pasly. Dense cataract since birth.	EEG background slow for age.	CT scan: right frontoparietal low density lesion with ex vacuo dilation of that ventricle and smaller right hemisphere. Presumed right MCA infarct	Epilepsy secondary to cerebral insult

GABRG2 c.211A>G (chr5:161520937)	36 months	Initial left sided seizures. Followed a few months later by some GTC seizures. Evolved to include drop attacks and myoclonic jerks, Seizures a few times a week.	*Trialled ASMs:* Carbamazepine Clobazam Lamotrigine Levetiracetam Phenobarbital Sodium Valproate Topiramate Response to ASMs not well described in notes	Described as normal.	Non-verbal. Evidence for neuroregression: normal early development - speech steadily lost over a few years.	Left hemiplegia and VII palsy Bilateral cataracts removed through surgery.	EEG initially normal but background evolved to become slow.	CT and MRI scans: right parietal and occipital atrophy corresponding to right sided PCA infarct. Left cerebral hemisphere mild atrophy Features suggest acute or chronic vascular insult.	Possible Lennox Gastaut Syndrome secondary to vascular insult
GRIN2B c.3499G>A (chr12:13716673)	48 months	Initially generalised tonic-clonic seizures. Evolved into myoclonic seizures about a year later. Absences noted about 6 months later.	*Trialled ASMs:* Carbamazepine Clobazam Lamotrigine Phenytoin Sodium Valproate	Described as normal.	Moderate global developmental delay. Developmental delay began after onset of epilepsy and accelerated as seizure disorder evolved.	None described.	Symmetrical 4-5hz theta activity with intermittent 1-3Hz transients.	CT and MRI - global atrophy	Doose syndrome

		Astatic / drops noted a year after that. Seizures once a week							

NOTE All patients met the criteria for being described as refractory to pharmacological treatment. NCBI RefSeq transcript references: *SCN1A*: NM_006920.5; *GRIN2A*: NM_001134407.2; *GABRG2*:198903.2; *GRIN2B*: NM_000834.4.

3.3 Pharmacogenomic results

Two analyses were performed: a broad pharmacogenomic analysis involving 68 general ADME variants associated with a wide range of drug metabolism phenotypes, and a second, more specific analysis focused on eight variants tentatively associated with the metabolism of ASMs in non-African populations. Those eight variants were chosen based upon their association, described in literature, with ASMs that have been prescribed to local state patients, including those participating in this study.

After stratification of the study participants based on ancestry, only the group of Black African ancestry was of sufficient size (n = 27) for the Pearson Chi-squared test to be performed, as the groups of other ancestries were too small and would have produced inaccurate statistical results (Mixed Ancestry n = 11, Indian Ancestry n = 1, European Ancestry n = 1). All results below are thus limited to the participants of Black African ancestry.

3.3.1 Veridose® Core Panel analysis

Of the 68 variants under investigation, 38 could not be analysed for Hardy-Weinberg equilibrium, nor undergo the Chi-square test, as no variant allele was detected in the study population for each of those variants (Table 7).

Table 7 Variants (n=38) in the broad pharmacogenomic analysis which could not undergo statistical testing due to the distribution of genotypes.

Gene	Variant
CYP1A2	rs12720461
	rs56107638
	rs72547513
CYP2C9	rs1799853
	rs56165452
	rs28371686
	rs9332131
	rs9332239

	rs72558187
	rs1304490498
	rs72558190
CYP2C19	rs4986893
	rs28399504
	rs56337013
	rs72552267
	rs72558186
	rs41291556
CYP2D6	rs35742686
	rs5030655
	rs5030867
	rs5030865
	rs5030656
	rs201377835
	rs5030862
	rs774671100
	dup4125_4133
	rs72549353
	rs72549354
CYP3A4	rs55785340
	rs4987161
	rs35599367
CYP3A5	rs28365083
F2	rs1799963
F5	rs6025
OPRM1	rs1799971
PNPLA5	rs5764010
SLCO1B1	rs4149056
SULT4A1	rs763120

Of the 30 variants which could undergo statistical analysis, one variant (rs59421388 in *CYP2D6*), appeared to be trending towards significance, with a P-value of 0.05 for the Chi-square test. Two variants (rs1057910 in *CYP2C9* and rs7900194 in *CYP2C9*), could be subjected to the HWE analysis but not the Chi-square test due to the distribution of genotypes

in the reference population (Table 8). For almost all variants, the genotype distribution in this study group was not significantly different to the general African reference population used for the analyses.

Table 8 General pharmacogenomic (ADME) variant (n=30) genotype counts, Hardy-Weinberg equilibrium P-value, and Pearson's Chi-squared test P-Value

Gene	Variant	Genotype counts	Genotype counts for African control populations	HWE P-value	Pearson's Chi-squared Test P-value
ABCB1	rs1045642	n = 28 G/G = 22 A/G = 6 A/A = 0	n = 661 G/G = 478 A/G = 168 A/A = 15	0.53	0.62
APOE	rs429358	n = 24 T/T = 15 C/T = 6 C/C = 3	n = 661 T/T = 356 C/T = 256 C/C = 49	0.10	0.32
	rs7412	n = 27 C/C = 20 C/T = 6 T/T = 1	n = 661 C/C = 530 C/T = 126 T/T = 5	0.53	0.24
COMT	rs4680	n = 28 G/G = 10 G/A = 15 A/A = 3	n = 661 G/G = 348 G/A = 255 A/A = 58	0.45	0.21
CYP1A2	rs2069514	n = 28 G/G = 18 G/A = 9 A/A = 1	n = 661 G/G = 307 G/A = 294 A/A = 60	0.92	0.16
	rs762551	n = 28 A/A = 11 C/A = 12 C/C = 5	n = 661 A/A = 213 C/A = 317 C/C = 131	0.59	0.74

CYP2B6	rs28399499	n = 28 T/T = 24 C/T = 4 C/C = 0	n = 661 T/T = 560 C/T = 93 C/C = 8	0.68	0.84
	rs3745274	n = 28 G/G = 12 G/T = 13 T/T = 3	n = 661 G/G = 259 G/T = 309 T/T = 93	0.85	0.86
CYP2C19	rs4244285	n = 28 G/G = 16 A/G = 12 A/A = 0	n = 661 G/G = 459 A/G = 179 A/A = 23	0.15	0.14
	rs12248560	n = 28 C/C = 19 C/T = 8 T/T = 1	n = 632 C/C = 384 C/T = 243 T/T = 5	0.89	0.20
CYP2C9	rs1057910	n = 28 A/A = 26 C/A = 2 C/C = 0	n = 661 A/A = 658 C/A = 3 C/C = 0	0.84	N/A
	rs7900194	n = 28 G/G = 23 G/A = 5 A/A = 0	n = 661 G/G = 591 G/A = 70 A/A = 0	0.60	N/A
	rs28371685	n = 28 C/C = 27 C/T = 1 T/T = 0	n = 661 C/C = 630 C/T = 30 T/T = 1	0.92	0.95
CYP2D6	rs16947	n = 28 G/G = 5 G/A = 11 A/A = 12	n = 661 G/G = 143 G/A = 304 A/A = 214	0.39	0.51
	rs1135840	n = 19 G/G = 8 C/G = 8 C/C = 3	n = 661 G/G = 318 C/G = 258 C/C = 85	0.68	0.86

	rs3892097	n = 27 C/C = 26 C/T = 1 T/T = 0	n = 661 C/C = 589 C/T = 64 T/T = 8	0.92	0.48
	rs1065852	n = 28 G/G = 24 A/G = 4 A/A = 0	n = 661 G/G = 529 A/G = 115 A/A = 17	0.68	0.61
	rs28371706	n = 12 G/G = 7 A/G = 5 A/A = 0	n = 661 G/G = 420 A/G = 194 A/A = 47	0.36	0.47
	rs59421388	n = 28 C/C = 19 T/C = 7 T/T = 2	n = 661 C/C = 529 T/C = 122 T/T = 10	0.27	**0.05**
	rs28371725	n = 28 C/C = 25 T/C = 3 T/T = 0	n = 661 C/C = 639 T/C = 20 T/T = 2	0.76	0.08
CYP3A5	rs776746	n = 28 T/T = 23 T/C = 5 C/C = 0	n = 661 T/T = 447 T/C = 190 C/C = 24	0.60	0.23
	rs41303343	n = 28 D/D = 18 D/A = 10 A/A = 0	n = 661 D/D = 516 D/A = 134 A/A = 11	0.25	0.12
	rs10264272	n = 28 C/C = 20 C/T = 7 T/T = 1	n = 661 C/C = 474 C/T = 170 T/T = 17	0.70	0.95
DRD2	rs1800497	n = 28 A/A = 5 A/G = 14 G/G = 9	n = 661 A/A = 91 A/G = 327 G/G = 243	0.91	0.95

GLP1R	rs1042044	n = 28 A/A = 6 C/A = 11 C/C = 10 C/T = 1	n = 661 A/A = 111 C/A = 322 C/C = 228 C/T = 0	0.38	0.66
	rs2300615	n = 28 T/T = 22 G/T = 5 G/G = 1	n = 661 T/T = 575 G/T = 83 G/G = 3	0.33	0.07
	rs6923761	n = 28 C/C = 27 G/C = 1 G/G = 0	n = 661 C/C = 643 G/C = 17 G/G = 1	0.92	0.93
MTHFR	rs1801131	n = 28 T/T = 21 G/T = 5 G/G = 2	n = 661 T/T = 478 G/T = 166 G/G = 17	0.07	0.27
	rs1801133	n = 28 G/G = 25 G/A = 3 A/A = 0	n = 661 G/G = 549 G/A = 105 A/A = 7	0.76	0.64
VKORC1	rs9923231	n = 28 C/C = 24 C/T = 4 T/T = 0	n = 661 C/C = 592 C/T = 66 T/T = 3	0.68	0.72

NOTE The Pearson's Chi-square test was performed to compare genotype distributions in the study participants of Black African ancestry to genotype distributions recorded for the general African population in the online database Ensembl.

3.3.2 Custom anti-seizure medication-specific mass array analysis

Analysis of the custom-designed mass array for ASM-associated SNVs, revealed two variants (*HTR2C* rs1414334 and *SCN1A* rs3812718) deviating from Hardy Weinberg equilibrium. One variant, rs1051740 in *EPHX1*, which is associated with the metabolism of carbamazepine in East Asian and European populations, appears to be significantly differently distributed in this

research group in comparison with the data available for the general African population, with a P-value of 0.02. Two other variants appear to be trending towards significance – rs2234922 in *EPHX1* (associated with carbamazepine metabolism in East Asian populations) with a P-value of 0.08, and rs3812718 in *SCN1A* (associated with the metabolism of phenytoin in East Asian and European populations) with a P-value of 0.05 (Table 9).

Table 9 ASM-specific variant genotype counts, Hardy-Weinberg equilibrium P-value, and Pearson's Chi-squared test P-Value

Gene	Variant	Genotype Count	Genotype Count in African Control Population	HWE P-value	Pearson's Chi-squared Test P-value
EPHX1	rs1051740	n = 28 T/T = 16 T/C = 9 C/C = 3	n = 661 T/T = 492 T/C = 152 C/C = 17	0.34	**0.02**
HTR2C	rs1414334	n = 26 G/G = 15 G/C = 3 C/C = 8	n = 661 G/G = 254 G/C = 176 C/C = 231	**0.001**	0.10
DRD2	rs1799978	n = 28 T/T = 16 T/C = 11 C/C = 1	n = 661 T/T = 459 T/C = 175 C/C = 27	0.59	0.33
ANKK1	rs1800497	n = 28 G/G = 9 G/A = 14 A/A = 5	n = 661 G/G = 243 G/A = 327 A/A = 91	0.91	0.79
UGT1A4	rs2011425	n = 28 T/T = 27 G/T = 1 G/G = 0	n = 661 T/T = 562 G/T = 96 G/G = 3	0.92	0.24
EPHX1	rs2234922	n = 27 A/A = 17	n = 661 A/A = 278	0.87	0.08

		G/A = 9 G/G = 1	G/A = 299 G/G = 84		
SCN1A	rs3812718	n = 24 C/C = 16 C/T = 5 T/T = 3	n = 661 C/C = 287 C/T = 299 T/T = 75	**0.04**	0.05
NA – Intergenic variant	rs489693	n = 28 C/C = 8 C/A = 16 A/A = 4	n = 661 C/C = 179 C/A = 298 A/A = 184	0.37	0.26

NOTE The Pearson's Chi-square test was performed to compare genotype distributions in the study participants of Black African ancestry to genotype distributions recorded for the general African population in the online database Ensembl.

Chapter 4: Discussion

4.1 Overview

Forty paediatric patients affected with epilepsy, and their available parents, were recruited into this project under the auspices of the parent project, “Precision management of epilepsy in South African children”. Thirty eight of the 40 participants donated saliva samples for DNA extraction, the remaining two donated blood samples from which DNA was extracted. A gene panel consisting of 78 genes was designed for targeted NGS on the Thermo Fisher Scientific Ion Torrent platform to facilitate the introduction of cost effective genetics-based diagnostic testing for paediatric epilepsy. All 40 participants were successfully sequenced using the gene panel. Sequencing data underwent variant prioritisation, followed by variant confirmation by direct cycle sequencing, and segregation analysis using available parental DNA. Three pathogenic variants were identified: two in *SCN1A* and one in *GRIN2A*, with a pickup rate of 7.5% (3/40). The findings aligned with the electroclinical features in each patient. Two VUSs were identified in *GABRG2* and *GRIN2B*. The findings were compatible with the electroclinical features in each patient.

Participant DNA also underwent pharmacogenomic analyses with both a commercial generalised pharmacogenomic panel and a custom-designed SNP array. The generalised pharmacogenomic analysis revealed one variant, rs1051740 in *EPHX1*, which appeared to be statistically significantly differently distributed in the study group compared to the reference population that was used as a proxy for the general African population (P = 0.02). Three other variants analysed with the custom-designed array, rs2234922 in *EPHX1,* rs3812718 in *SCN1A*, and rs59421388 in *CYP2D6,* appeared to be trending towards significance. Both, the primary and secondary aims of this project, and their objectives, as outlined in Chapter 1, were thus achieved.

4.2 Findings

The study group was phenotypically heterogenous, owing to the broad inclusion criteria of the parent project (phenotypic details in Appendix E), with all participants falling under the broad description of clinically defined complex epilepsy with drug-resistance. The reasons behind the age criterion of the study (inclusion criteria being 4-16 years of age at time of recruitment)

included the practical consideration that children had to be old enough to wear the watch used for biometric monitoring in the parent study. Additionally, collection of saliva samples is minimally invasive and traumatic for children above the age of 4, in comparison to the drawing of blood.

Collection of saliva with the Oragene Collection kit specifically has additional practical advantages over collection of blood samples – the Oragene kit allows saliva samples to be stored for a substantial length of time (up to several months) at room temperature until DNA is extracted, in contrast to blood samples collected through normal phlebotomy, which must be refrigerated and undergo DNA extraction timeously[106]. This practical advantage was useful, as it enabled sampling at different sites, with less risk of sample degradation upon transport and prolonged delay prior to extraction. Sampling could also be performed by the participants themselves with the aid of instructions, rather than requiring trained personnel to perform phlebotomy. The major disadvantage of the Oragene Kit was its expense in comparison to other DNA extraction methods (either manual or automated). The advantages of sample storage without degradation, which may justify the kit cost, is relevant mainly in the context of research, but is not so relevant in the context of routine diagnostic/clinical practice. In this study, the participant samples were collected during a process of recruitment into the parent study that neared completion before the beginning of this sub-study – the ability to store samples was thus important. In contrast, in routine diagnostics, samples would be processed much more rapidly, and so a long shelf-life prior to DNA extraction becomes potentially redundant.

A disadvantage of the use of saliva samples (in general, not only those collected using the Oragene kit) is that the DNA extracted from saliva tends to be of poorer purity than DNA extracted from blood. Following extraction from saliva samples, a wide range of DNA concentrations were produced (3.19ng/uL to 2644.62nguL). A lower DNA concentration is not necessarily a barrier to successful NGS, given the sensitivity of the NGS processes, and the low starting DNA concentration required (1ng/uL) to begin NGS. Of greater concern were the poor 280/260 and 260/230 quality ratios. The 280/260 ratio is an absorbance ratio measured by the NanoDrop Spectrophotometer, which is used to interrogate the purity of a sample. A ratio of 1.8 is considered an indicator of “pure” DNA, while deviation of the ratio may indicate the presence of contaminants that absorb at or near to 280nm, such as proteins. The 280/260 ratio across the samples ranged between 1.26 - 2.38, though the majority of samples (61/90,

67.78%) clustered around a ratio of 1.8 (that is, deviated by 0.1 or less from the ideal ratio), and so was deemed acceptable for most patients.

The 260/230 ratio was of greater concern and ranged more widely, from as low as 0.30 to as high as 2.50. The 260/230 ratio is a secondary measure of purity, where a reading of 2.0-2.2 is taken as an indicator of a good quality sample. Abnormal ratios may indicate the presence of contaminants that absorb at 230nm, such as EDTA, phenol, and carbohydrates[107]. No sample produced a 260/230 ratio in the expected range of 2.0-2.2, with all but one sample producing ratios below that range (Appendix F). This indicates a high level of contamination in the samples by contaminants which absorb at 230nm, potentially organic compounds or carbohydrates. The reason for this may be the presence of substances in the human mouth, such as non-human DNA from bacteria and fungi, food particles, toothpaste, lipstick, and so on. However, this contamination may be reduced in the future by careful attention on the part of researchers and participants to collection instructions. Such contamination is a trade-off for the non-invasive collection of saliva samples, and although studies have found DNA from saliva to be consistently of lower quantity and quality than DNA from blood, DNA from saliva may often be adequate for traditionally robust genotyping involving fragment analysis or cycle sequencing[106]. Such genotyping processes would require more total DNA than the NGS protocol used in this project, but may be less sensitive to the presence of contaminants. Most current NGS protocols, however, involve diluting DNA to very low concentrations (1 - 10ng/uL), which also dilutes any present contaminants to levels that may no longer compromise the success of the downstream applications. Several DNA samples showed poor amplification with the RNase P assay and were subjected to clean-up using PCR inhibitor removal kits and DNA clean-and-concentrate kits. Subsequently, those samples showed marginally improved amplification in the RNase P assay, but at the cost of a dramatic decrease in DNA concentration. Due to that decrease in DNA concentration, no further samples were processed through the clean-up step, as decreasing the DNA concentration of the samples, whilst not impeding NGS processes, could impede or prohibit the success of other downstream applications like cycle sequencing, genotyping, or chromosomal arrays.

After all samples had undergone NGS, the average depth of coverage observed for those samples ranged from 206.2 to 486.9, with an overall average of 363.51. This is comfortably above the desired depth of coverage for diagnostic purposes, validating the use of DNA extracted from saliva for NGS applications. However, subsequent pharmacogenomic analysis and chromosomal microarray analysis (CMA) (done outside of the scope of this project) were

problematic, as some of the DNA samples did not meet the quality control criteria set by the technical service providers performing those analyses, even after additional clean-up. The pharmacogenomic analyses were performed for all patient samples, but several patient samples were not processed through CMA. Thus, when considering future sample collection, one of the considerations when choosing saliva should be the type of testing which is planned, and the DNA quality required for that testing. It is also worth noting that two patients had DNA extracted from blood samples, and those samples, despite one being very low in concentration, encountered none of the same issues as the samples extracted from saliva, and were successfully and easily processed through all analyses and laboratory procedures. In routine local practice, it is thus likely that blood samples will be used more often than saliva samples to provide DNA for testing.

The choice of the diagnostic approach, that is, which NGS protocol to develop for translation, required the consideration of a range of clinical, technical, and practical factors. Affordability, ease and speed of analysis, data storage capacity, and personnel and facility requirements were all important considerations for the local setting. Regarding the potential selection of ES as the NGS protocol, several aspects were considered. ES produces a lower depth of coverage on average than gene panels, and increasing the depth of coverage for ES raises its cost. ES also has a high rate of exonic or allele drop-out (higher than the gene panel alternative), where certain regions are not sequenced, which may result in false negatives. This is due to the enormous breadth of coverage required by ES, in that it must capture the entirety of the exome, to produce data for approximately 20 000 genes. A major strength of ES in diagnostics (and research) is the option for retrospective re-analysis of sequencing data, as new knowledge becomes available, and new clinically relevant genes and variants are described.

Whilst ES (and even GS) has been adopted as the first-tier analysis for genetic epilepsy in many laboratories in HICs[46], the cost of ES is still too high for routine testing in the South African public health sector. However, if accessible, it may be more economical to sequence a single exome in cases where a phenotypic presentation is particularly complex, rather than sequence multiple panels or perform multiple other diagnostic tests. ES requires extensive data storage facilities, and sophisticated data analysis approaches to cope with sequencing data for over 20 000 genes. Moreover, ES requires highly trained personnel and machines with specific capabilities. An additional complication of performing diagnostic ES in a genomically underrepresented SSA patient population is the great genetic variation of the SSA

population, which is not reflected in international browsers and databases. Compared to other ancestries, individuals of African descent have a significantly higher number of previously uncharacterised variants, producing larger variant call files, with high numbers of VUSs, especially if parental samples are not available to establish *de novo* occurrence.

Finally, the matter of incidental findings with ES must be considered. Incidental findings refer to the detection of pathogenic variants which are not related to the initial reason for testing, producing ethical dilemmas related to result delivery and clinical management[46]. Such ethical dilemmas may be exacerbated in contexts where there is a lack of easy access to medical geneticists and genetic counselling services. The handling of incidental findings must be carefully considered, and protocols are required to be drawn up for individual laboratories, based on international recommendations. Incidental findings may arise in ES due to all genes being sequenced, not just those related to the disorder under investigation.

Gene panels, as the alternative approach, regularly produce a depth of coverage greater than required for diagnostic purposes, allowing for confident and reliable detection of putative variants. In terms of depth of coverage, the choice of NGS protocol is thus weighted in favour of gene panels. A limitation of gene panels is that re-analysis is not possible to the extent it is when performing ED – gene panels are limited by their design, and may only be reanalysed in light of new data about the particular genes included in the panel. However, a panel may be continually updated and reordered from the manufacturer, a process which could be streamlined through active professional relationships between laboratories and commercial entities.

Panels, in contrast to ES and GS, produce less data, and both analyses and data storage are more accessible and easily implemented. Panels are also an accessible technique which has been highly optimised and streamlined by manufacturers. Extensive training is not required to perform the sequencing process nor the analysis, and thus the use of panels is more easily integrated into pre-existing services. Panels also generally do not produce incidental findings, given that all the genes which are sequenced are specific to the disease or disorder under investigation. The popular approach in the HICs is that of the so called “virtual panels”, where exome data is analysed by targeting bioinformatic pipelines at specific sets of genes. This negates the issues of incidental findings, and the ES data is available for future re-

assessment. However, this pragmatic option cannot be implemented locally at this time, as ES is not currently an affordable testing modality in the South African state healthcare setting.

Thus, taking into account the above, gene panels were decided to be the most appropriate NGS protocol for the South African public sector diagnostic context. Although this approach may be out-of-step with the current recommendations and practices in HICs, where ES is offered as first-tier analysis (and GS is entering the diagnostic setting), the implementation of a diagnostic protocol in South Africa requires adaptation to fit the local resource-constrained setting[46,96]. Implementation of ES in Africa as a routine first choice diagnostic test will ultimately occur, but this is likely to be a gradual process, whereas there is an immediate need for an NGS-based test for genetic paediatric epilepsy. The choice of gene panels as the diagnostic approach will have the greatest accessibility and utility, and thus will have the most immediate beneficial impact in the South African public healthcare sector.

During the process of selecting genes for inclusion in the panel, careful consideration was given to the clinical context in which the panel will be used. The majority of recruited patients had previously received working clinical diagnoses of DEEs, with drug-resistant seizures since infancy. This was relevant to this genetic study as the success of gene discovery and the rate of positive findings, with implications for treatment, have been highest in the neonatal/infantile-onset DEEs[46]. The panel design was thus aimed mainly at the monogenic DEEs, where genetic testing currently has the greatest clinical utility. In African patients with a DEE of an assumed genetic cause, most cases occur without any known family history, and are assumed to arise from *de novo* pathogenic variants in autosomal dominant genes. In other regions of the world, such as the HICs, this has been observed to be the case, with the majority of DEEs being observed to occur due to such *de novo* variants in AD genes. It is thus reasonable to assume that the same AD genes observed to be causing DEEs in HICs are relevant to the DEEs in African patients, especially those genes known to perform important functions in development and neurological functioning[108]. It is thus reasonable to include in this panel those AD genes identified as being causative of epilepsy in other, non-African populations.

However, when it comes to recessively inherited epilepsies, there is a possibility that the panel was not picking up causative variants in the AR genes in the panel (15 of 78 genes) because these genes, identified in non-African populations, are in fact less relevant to African populations and variants in other, as yet unknown AR genes are more prevalent in Africa for

cases of recessively inherited epilepsy. It should thus be emphasised that although this panel has successfully identified epilepsy-causing variants in the patients in this study, once translated, it should be continuously curated based on its performance, and updated to keep it up to date with new research performed in Africa and elsewhere in the world.

Another consideration for gene selection for the panel was whether the underlying aetiology of an epilepsy is identifiable by other means, such as metabolic testing or MRI, and whether genetic variant detection is necessary to inform treatment. This, however, was not a *de facto* criterion for the exclusion of genes from the panel, and genes were considered in a case-by-case manner. For example, *ALDH7A1*, which causes pyridoxine-deficient epilepsy, was included on the panel, even though pyridoxine-deficient epilepsy is detectable by metabolic testing. Limited metabolic testing is available in the South African state healthcare service and is usually requested in patients whose clinical features may suggest metabolic disease. Epilepsies with an underlying metabolic cause present with a range of symptoms, not all of which are considered typical presentations[68,109]. Thus, the inclusion of *ALDH7A1* in the gene panel would be to enable rapid identification of atypically presenting cases, which is highly important considering the high actionability of the metabolic epilepsies and the benefits of early treatment. Further examples are genes causing disorders of cortical malformation, which may be diagnosed by imaging and treated by surgery[110]. Such genes were considered carefully regarding their inclusion. For example, the *AKT3* gene, which is implicated in hemimegaloencephaly, was not included on this panel, as hemimegaloencephaly is diagnosable by MRI and managed largely by surgery[111].

In contrast, the *TSC1* and *TSC2* genes were included in the panel, even though Tuberous Sclerosis Complex (TSC) is characterised by focal cortical malformations detectable on MRI and can usually be clinically diagnosed. In many cases though, TSC initially presents with epileptic spasms, which is a clinical entity among multiple DEEs. In 2010, the mTOR antagonist rapamycin received FDA approval for treatment in TSC patients who cannot be treated with surgery[112]. Patients with *DEPDC5*-associated familial focal epilepsy with variable foci may also exhibit cortical malformations. The *DEPDC5*-encoded protein is also involved in the mTOR pathway, implying possible further use for mTOR inhibitors such as rapamycin or everolimus in treatment of those cases[113]. Thus, inclusion of these genes in this panel, despite their association with cortical malformations diagnosable by imaging, was motivated by the potential clinical actionability of positive genetic findings. It was also recognised that a germline DNA testing approach may not be appropriate for some genes and disorders. For example,

the *GNAQ* gene implicated in Sturge-Weber syndrome was not included in the panel as pathogenic variants in *GNAQ* occur after conception, and are therefore by definition mosaic, with a very low risk of recurrence in the family. This somatic mosaicism makes genetic diagnosis of Sturge-Weber syndrome difficult and less effective than diagnosis by clinical examination and imaging[114]. Once genes had been selected for inclusion in the panel, they were cross-referenced to the ClinGen Expert Panel resource for epilepsy gene curation (https://clinicalgenome.org/affiliation/40005/), an important resource on the most up-to-date classifications of causative or associated epilepsy genes. Not all the genes included in that resource are relevant to this panel (as this panel was designed to be DEE-focused, whilst the ClinGen epilepsy database is much broader in scope), but this was a necessary step to double check the appropriateness of including the selected genes as genes of definitive clinical validity.

Of the 78 genes selected for the panel, 73 were available as "On Demand" genes from Thermo Fisher Scientific. Amplicons of the On Demand genes are optimised and pre-tested by the company for verified performance and near 100% capture. Five genes (*ARX*, *FOXG1*, *GRIN2D*, *NR2F1*, and *PURA*) were not available as On Demand genes and had to be ordered as a separate "spike-in" panel. Amplicons of the spike-in genes are not pre-tested by the manufacturer and the *in-silico* coverage predicted approximately 70% coverage across that entire panel. As a rule, gaps in coverage are related to the difficulty in amplifying a region that may be GC-rich or comprised of repetitive sequences (e.g., homopolymer stretches)[115]. Filling these gaps, if they are observed in clinically relevant regions deemed important for the overall test performance, may involve ancillary assays, such as cycle sequencing.

Despite the *in silico* prediction of missing sequences in the spike-in genes, those genes were still included due to their clinical relevance, and the fact that the prediction was an *in silico* prediction, whereas in practice those regions may still have been covered. No missing sequence was observed in *NR2F1*. Missing sequences were observed in four genes of the spike-in panel, consistently across all samples, in in *ARX* exon 2, *FOXG1* exon 1, *GRIN2D* exons 2, 3, 4, and 13, and in *PURA* exon 1. The context of the missing sequence regions was visually inspected in IGV, all of which mainly consisted of long sections of GC rich regions and stretches of homopolymers (described in Chapter 3), likely to cause or contribute to the sequence drop-out. Going forward, consideration should thus be given to whether or not, these gaps in coverage warrant design of the so-called "filler tests" (e.g., PCR and/or cycle

sequencing) or if the limitation should simply be a caveat explained in the diagnostic test report.

Regarding *ARX,* the most commonly occurring pathogenic variants have been observed to occur in exon 2[116]. In one case, the loss of exon 5 was reported to be due to a splice-site variant in intron 4 rather than a variant in the exon itself[117]. It may thus be inferred from published data that the missing sequence of exon 2 is particularly problematic and may affect the detection of causative variants in *ARX*, as the commonly reported variants found in exon 2 would not be detected. It may thus be necessary to design ancillary assays for the *ARX* gene, particularly in cases where patients are presenting with phenotypic features associated with pathogenic variants in *ARX.* Considering *FOXG1,* that gene consists of only one exon, so the missing sequence in that exon is also problematic. However, *FOXG1*-syndromes have been reported to have stereotypical clinical presentations resembling Rett syndrome, and so specific testing for variants in *FOXG1* by means other than the panel, such as cycle sequencing, could be directed by phenotype[118].

GRIN2D is a large gene of 13 exons, and despite the drop-out of exons 2-4, a large proportion of the gene (nine exons) is still covered. Additionally, actionable variants with precision medicine responses have been reported in *GRIN2D* and so on balance its inclusion in the panel remains important, but testing of *GRIN2D* may have to be supplemented with other tests, or the amplicon design may have to be altered to attempt to achieve greater coverage[46,119]. Finally, the missing sequence in *PURA* presents a complicated scenario. Studies on *PURA* have found pathogenic variants in the gene, though without any clear genotype-phenotype correlations, complicating a phenotype directed focus on *PURA* testing[120]. Additionally, specific treatment recommendations for *PURA* variants are not yet developed[120]. However, *PURA* consists of a singular exon, and 871 of 979 base pairs are covered (89%), and so its continued inclusion in the panel may be justifiable on that fact. As technical advances continue, recommendations for dealing with missing sequence regions include improving the detection of missing sequence, while trying to minimise such occurrences[121]. As this panel is translated for diagnostic use, a database of regions which tend not be covered by sequencing should be kept to help highlight those regions and decrease the likelihood of a false negative due to missing sequence.

During sequencing, different concentrations of prepared library were considered for use in the templating process, in order to find an appropriate balance between ISP loading and polyclonality. For the first sequencing run, a library concentration of 30pM was used, which produced 80% usable reads. In the second sequencing run, a concentration of 35pM was used, which produced 64% usable reads (Chapter 3, Figure 1A and 1B). Despite the lower concentration producing a higher percentage of usable reads, its resulting ISP loading was substantially lower (Chapter 3, Figure 1A). As ISP loading is a key step in the amplification of DNA during sequencing, the lower loading resulted in fewer total reads than the higher concentration (a difference of just over half a million between sequencing runs 1 and 2). As such, although the higher concentration of 35pM produced a lower percentage of usable reads, it produced more reads overall (and the percentage of usable reads was still deemed acceptable). It was therefore decided that the higher concentration would be used for the remaining sequencing runs. A greater number of total reads was deemed important as it translates into a higher overall DOC, which is a key quality indicator for diagnostics.

Following variant prioritisation, confirmation, segregation analysis, and classification of pathogenicity according to ACMG criteria[122], five variants of interest were identified in five probands: three pathogenic variants (two in *SCN1A* and one in *GRIN2A*), and two VUSs in *GABRG2* and *GRIN2B* (Chapter 3, Table 5). An important final step in the variant classification process was consultation with a multidisciplinary clinical team.

Two pathogenic variants were detected in *SCN1A*. The first, a c.4444-1C>T splice-site variant, is a LoF variant, and has been described before (ClinVar ID: 530456; rs1553521567). The variant was detected in a patient clinically diagnosed with early onset epileptic encephalopathy prior to genetic testing, who presented with tonic clonic seizures at the age of two months (Chapter 3, Table 6). The identification of this has treatment implication for the patient – channel blocker medications should be avoided[46]. The second *SCN1A* variant, a nonsense c.664C>T variant, was identified in a patient clinically diagnosed with Dravet Syndrome, who presented with focal seizures with secondary generalised seizures at the age of five months (Chapter 3, Table 6). This variant is also predicted to be a truncating LoF variant, previously identified and published by several investigators (ClinVar ID: 12889; rs121918624). Channel blocker medications should also be avoided for this patient[46]. Additionally, both these patients with pathogenic variants in *SCN1A* had been included in another internal genetic study at UCT, with sequencing performed in another laboratory using Illumina sequencing technology, and the results obtained across both studies were the same. This may be seen as additional

(internal) quality control step, demonstrating the panel's ability to pick up on variants if they are present and replicate the findings of other diagnostic tests.

The third pathogenic variant was identified as *GRIN2A* c.2191G>A, which has been described before (ClinVar ID: 205657; rs796052549), and is described as a LoF variant in online databases. The variant was identified in a patient who presented with left-sided clonic jerks at 72 months of age, and who also had a dense cataract in the left eye and left-sided hemiplegia. The cataract and hemiplegia are considered possible co-morbidities. The initial working diagnosis for this patient was epilepsy secondary to a central nervous system insult, but the identification of a pathogenic variant in this patient suggests a genetic component to the development of epilepsy symptoms and associated co-morbidities, making this patient a complicated case of combined genetic and environmental causes of epilepsy and developmental delay. The treatment implications for this variant include potential treatment with serine, though currently the evidence for this treatment is from retrospective clinical case studies, and not clinical trials[46].

Regarding the two identified VUSs, *GABRG2* c.211A>G has not been described before, while *GRIN2B* c.3499G>A has (ClinVar ID: 245684; rs1042339). *GRIN2B* c.3499G>A has a "one star" rating in ClinVar, essentially indicating conflicting interpretations of pathogenicity across different studies. The disease mechanism of action is not described for either of these variants. No specific precision medicine treatments are described for *GABRG2*. Specific drugs may be effective in the case of *GRIN2B,* but which drug to use depends on the whether the variant is GOF or LOF, which is not currently described by any functional studies for the c.3499G>A variant[46]. Reanalysis and reclassification of the VUSs may be possible if more data on patients, the variant, or the gene become available over time. Changing or directing clinical treatment based on the identification of a VUS is not recommended, but clinicians may use their discretion when treating the patients while understanding the limitations of these results.

Demonstrating *d*e *novo* occurrence by testing the patients' parents is an essential part of variant classification. However, the study group comprised mainly of mother-child duos, as recruitment took place in "real-time", during clinic visits. Owing to the socio-economic realities of South Africa, both parents of an affected child may not be available to attend clinics and provide DNA samples[4]. Being able to test both parents is particularly important in epilepsy diagnostics, given that most (though not all) causative variants in the severe paediatric

epilepsy phenotypes occur *de novo* in AD genes. Access to patient-parent trios would likely raise the panel's diagnostic yield and/or help to resolve the VUSs, however, the absence of DNA from both parents does not preclude identification of pathogenic variants. In some cases, requesting the second parent's sample may be possible as part of genetic counselling. In the absence of DNA from both parents, segregation analysis could not be performed on variants in two patients in this study, (*SCN1A* c.4444-1C>T and *GRIN2A* c.2191G>A), but the variants were still classified as pathogenic, based on the functional effect (truncation and predicted LoF), absence from population databases, recurrence, and correlation with phenotype, in keeping with the ACMG criteria[122].

Our panel thus demonstrated a diagnostic yield of 3/40(8%). This is on the lower end of diagnostic yields reported in literature[123], but this is likely due to the small number of study participants selected for the parent study using very broad clinical criteria, a number of whom were thought to have infectious or structural aetiologies or contributors. Not excluding patients due to suspected or attributed acquired causes is important as dual diagnoses and incorrectly ascribed aetiology to infectious causes should be a consideration, and not preclude access to epilepsy genetic testing in appropriately selected patients. This approach was taken specifically into account when opting to sequence all patients recruited for the parent study, and the likely dual diagnosis in Patient 31 of both the pathogenic *GRIN2A* c.2191G>A variant validates this approach.

As more patients, specifically selected for genetic testing by clinicians, undergo sequencing, the diagnostic yield is likely to increase. The yield in this group could increase to 5/40 (12.5%), should the VUSs be upgraded to likely pathogenic or pathogenic variants in the future. This may be possible with access to parental DNA for segregation studies to prove *de novo* occurrence, or new functional studies showing evidence of the deleterious effect of the variants. It is equally possible however, that the variants are benign and common among the local population, and their absence from international databases is simply due to the current underrepresentation of genomic variation in individuals of Sub-Saharan African descent in those databases. It is worth noting that the inclusion criteria of a minimum age of 4 years may have also affected the diagnostic yield, as it excluded younger children, a section of the population more likely to have monogenic epilepsy detectable by the panel. This inclusion criterion was for practical considerations for the parent study under which recruitment occurred, and will not factor in routine clinical diagnostic testing. Finally, minor alterations to the filtering process for identifying variants of interest could also potentially improve the

diagnostic yield – filtering for variants in AD genes required variants to be absent in the general population, but the use of a very low minor allele frequency (such as 1-2%) may be more appropriate to accommodate for genes, variants, and phenotypes associated with which are not totally penetrant.

The primary aim of this project was to develop a NGS gene panel for detection of pathogenic variants in patients with early-onset, drug-resistant genetic epilepsy. Additional diagnostic protocols, such as CMA, to investigate genomic copy number variants, whilst relevant, fell outside of the scope of this project. CNV analysis can be performed using NGS panel data, especially for exon-level CNVs, but in the diagnostic sphere, the inferred copy number changes are generally confirmed with another method, such as MLPA or CMA[124]. The Ion Torrent sequencing platform is able to detect CNVs from NGS data, but this analysis requires the generation of what is termed a "baseline". This baseline requires training data of sequencing data from individuals with epilepsy, but epilepsy without a CNV cause – that is, it requires sequencing data from cases with confirmed causative pathogenic SNVs in order to detect cases of CNV-caused epilepsy. This study's sample size was not considered sufficient to create the baseline required for computational analysis of NGS data to detect CNVs, but as that sample size increases, as the panel is used in the diagnostic laboratory, such computational analysis should become possible. However, genome wide CNV analysis using CMA is a relevant investigation in the DEEs, especially in the phenotypes involving pre-existing features of intellectual or developmental disabilities[125].

For diagnostic completeness, those probands with no informative findings with the NGS panel, were selected for CMA by a clinical geneticist, as part of the study's parent project. Twenty patients were referred to the local National Health Laboratory Service (NHLS) for diagnostic CMA testing using the Affymetrix CytoScan® 750K array (Thermo Fisher Scientific). Only one of the twenty patients had a positive finding (a deletion in the Angelman Syndrome/Prader Willi Syndrome region of Chromosome 15q11). This diagnostic rate of 5% is within the range of what is expected for CMA diagnostic yields, however, it is only a single result from a small sized study group, and so may not necessarily reflect the true rate of CNV-caused epilepsy in the population. More patients would have to be tested to confirm that diagnostic rate, and the continued application of CMA analysis in tandem with the NGS protocol will be necessary in clinical practice for diagnostic completeness[46].

The secondary aim of this project was to conduct a preliminary exploration of the pharmacogenetics of ASMs and drug-resistant epilepsy in the study participants, for exposure to pharmacogenomic analysis and to provide direction for future pharmacogenomic research into drug-resistant epilepsy in South Africa. This preliminary investigation took a two-pronged approach. The first approach utilised a generalised pharmacogenetics panel (the VeriDose® Core Panel designed by Agena Bioscience (San Diego, California, USA)) of ADME variants previously associated with a collection of different drug metabolism phenotypes, but not previously associated with ASM metabolism, to investigate the possibility of these variants influencing drug-resistance in the study group. Data analysis was complicated by the small sample size and the observed genotype distributions, as 38 variants could not be analysed due to uniform genotype distribution across the study group. It is possible that variants were not detected because they simply do not exist in the South African population, or that the sample size of 40 patients was too small to detect variation.

Following this generalised analysis, one variant, rs28371725 in *CYP2D6,* was observed to be trending towards being significantly differently distributed to the reference population (P = 0.05). *CYP2D6* encodes a highly polymorphic enzyme implicated in the metabolism of antidepressants, antipsychotics, and analgesics, amongst other medications[126]. The interpretation of this result is uncertain, however, as this variant was assayed using the generalised panel, and it is uncertain which, if any, ASM metabolism is affected. Even if replicated in a larger sample set, the meaning of this result is uncertain due to the lack of specificity of the assay and the broad effects of the variant on the metabolism of a large number of different drugs.

The second approach involved the investigation of variants previously implicated in ASM metabolism in non-African populations, to investigate if those results might be replicated in this study group, to investigate their possible relevance to African populations. Of these variants, rs1051740 in *EPHX1*, which is implicated in the metabolism of carbamazepine in non-African populations, appeared to have a significantly different genotype distribution in the study group compared to the reference population used as a proxy for the general African population (P = 0.02). The small sample size of the study group necessitates investigation in a larger sample set before this result can be reported with confidence. Furthermore, functional characterisation is required to understand the variant's influence on carbamazepine metabolism. An additional variant in *EPHX1*, rs2234922, previously associated with carbamazepine metabolism in East Asian populations, appeared to be trending towards

significance with a P-value of 0.08. This, again, requires testing in a larger cohort. Potential identification of two carbamazepine-associated variants in *EPHX1* may point towards future pharmacogenomic research study with a specific focus on *EPHX1* and carbamazepine metabolism in African populations. Such research may provide translatable data and diagnostic protocols, such as *EPHX1* genotyping prior to carbamazepine prescription.

A variant in *SCN1A*, rs3812718, which has been observed to be associated with phenytoin metabolism in non-African populations, also appeared to be trending towards significance with a P-value of 0.05. Phenytoin is a sodium channel blocker, and thus a variant influencing its metabolism in *SCN1A,* a gene strongly associated with channelopathies, is potentially very interesting given that phenotypic link[127]. However, this variant is intronic (it is why this variant was investigated by array rather than using the sequencing data produced by the panel), so its mechanism of action would be difficult to predict, and again, the result would need to be replicated in study group of a larger sample size.

4.3 Challenges and Limitations

One key challenge in this study was the broad phenotypic inclusion criteria used during recruitment. This was not necessarily a challenge to the panel design itself, but has likely influenced the variant pick-up rate of the panel, which was designed mainly for patients with DEE. The true pick-up rate in the local clinical setting will be revealed over time, as the panel is used in the diagnostic sphere and patients with more defined phenotypes are referred for testing. The panel design was aimed mainly at the monogenic DEEs, as that is the patient population with the highest yield of clinically actionable and useful genetic findings. In view of the expense of genetic testing in the local setting, gatekeeping will be required to ensure the appropriateness of testing. Accurate phenotyping will be imperative, and in this regard, it is important to consider the realities of the clinical sector in South Africa.

A shortage of experienced paediatric neurologists entails difficulties in acquiring extensive and detailed phenotypic classifications of epilepsy in patients. Patients who should qualify for testing may “slip through the cracks”, especially at non-tertiary health centres, and especially if they have already been diagnosed with an assumed infectious or structural cause of their epilepsy. An aspect of successful implementation would thus require the development of

decision making tools to aid clinicians in deciding whether or not to refer a patient for genetic testing[96]. This is a particularly important aspect considering the beneficial effects a genetic diagnosis can have in directing treatment and improving prognoses, and considering the importance of choosing a correct diagnostic course in resource-constrained settings[96].

Another key challenge was the collection of DNA from both parents for each proband, and a low proportion in the study group of proband-parent trios. This is a common challenge in the South African context, and may be addressed over time and through communication with patients and health care providers to enable access to both parents. Though not ideal, the lack of parental DNA does not prevent the identification of a causative variant in every case, as described above. Two pathogenic variants were successfully identified without full segregation analysis being performed. The technical challenges also include the sequence drop-out in the spike-in panel. As commercial service and technology providers expand their list of optimised genes, those genes may well be optimised and this challenge circumnavigated. Until such time, this limitation may either be factored into the diagnostic test reports, or fill-in assays may be performed to cover the missing regions.

A key limitation of the pharmacogenetics portion of the project, was the small sample size of the study group which precluded any definitive conclusions based on the results, as was expected for this preliminary investigation. This challenge may be overcome in future studies through prolonged and collaborative recruitment drives. Whilst the small cohort size was not a factor for the NGS panel design and identification of disease-causing variants, it was a limitation of the pharmacogenetics portion of the study, as pharmacogenetics analysis investigates statistical associations, where strength of associations depends on sample size. An additional challenge for the pharmacogenetics analysis was the lack of population data for SSA populations, which limited the investigation and comparison of variant frequencies. One limitation of the study may be that the reference population used as a proxy for the general African population for these analyses included West and East African populations, which may be very distinct from Southern African populations, due to the innate high genetic diversity present across African populations.

4.4 Future Work/Directions

Enabling the support of a successful translation of the panel designed in this study into the diagnostic laboratory service for the DEEs is an important part of future work. This includes continuous curation of the panel to keep it up to date with the most current research and data. Establishing multidisciplinary team communication for feedback and discussion of findings will help establish the diagnosis and direct the care of patients. A key element will involve teaching and both clinical and genetic training of a range of professional groups and stakeholders, including doctors, nurses, and genetic counsellors[96]. More epilepsy research in local populations and research on other common epilepsy phenotypes will also be required to ensure the successful pick up of all possible causative variants in individuals of African descent. Easier access to databases of African variation will also be required in order to access allele frequencies to interpret novel (previously undescribed) variants. Finally, perhaps one of the most important aspects of successful translation will be ensuring investment from the South African state healthcare service, in order for appropriate funding to be allocated for the wide-scale provision of the panel as a routine service[96].

A key next step in pharmacogenomic research for epilepsy is recruitment of large cohorts for establishing statistically significant associations. Pharmacogenomic research on epilepsy requires complex and careful planning, and appropriate study design is essential for generating informative and useful results. The extensive phenotypic and genetic heterogeneity of epilepsy complicates recruitment and study design, and rare phenotypes and genotypes create difficulties in recruiting large cohorts. Establishing collaborations across multiple research centres will be an important step in ensuring adequate recruitment and study sample sizes. Refining the inclusion criteria for pharmacogenomic analysis, and focusing studies on drug resistance to a singular, specific ASMs, especially those widely used in the local state healthcare setting, may be one way of maximising the informativity of study results. Adverse drug reactions could also aid in directing research; extreme phenotypic reactions such as an ADR could act as inclusion criteria to ensure the selection of a research cohort in which some kind of abnormal ASM metabolism is present. However, this would have to be in the cases where the ADR was known to be a result of innate metabolism, rather than drug-drug interactions or other non-innate biological factors. Additionally, proteomic or transcriptomic research into particular types of drug resistance displayed, such as delayed or fluctuating drug resistance, could give insight at the non-genetic level into the metabolic mechanisms at play

in drug resistance, and thus allow for direction at the genetic level to new variants or genetic factors.

This study investigated variants already tentatively associated with ASM metabolism, but the investigation of what novel variants may be influencing ASM metabolism in African populations is also of great importance. That question would have to be addressed by substantial exploratory research including the use of a range of molecular techniques to identify and characterise new variants, such as ES, GS, and functional studies. Such research would require substantial investment of time and resources, but could ultimately produce results that would not just have benefit in Africa, but also across other regions, in combatting the substantial levels of ASM resistance observed globally. Additionally, this research is likely to become more accessible and easier to implement as genomic research in general increases in African populations and more data, especially sequencing data, becomes available to researchers.

4.5 Conclusion

The main outcome of this project was the successful development of an NGS-based diagnostic protocol for paediatric epilepsy using the gene panel approach, and the panel is now ready for translation to and validation in the local diagnostic testing laboratory. This will enable access to genetic testing for paediatric epilepsy for a far greater range of patients and families, and help enable research into genetic epilepsy in South Africa. Moreover, the genetic cause of epilepsy was identified in three study participants, with treatment implications in two (patients with variants in *SCN1A*). The pharmacogenomic analysis in this project did not reveal specific insights owing to the small study group size, but provided both exposure to pharmacogenomic analysis, and may potentially be used as a basis for further research into the pharmacogenetics of epilepsy in larger African cohorts, such as potentially investigating the influence of variants in *EPHX1* on the metabolism of carbamazepine.

References

1. Ngugi, A.K., Bottomley, C., Kleinschmidt, I., Sander, J.W., and Newton, C.R. (2010). Estimation of the burden of active and life-time epilepsy: A meta-analytic approach. Epilepsia *51*, 883–890. 10.1111/j.1528-1167.2009.02481.x.

2. Kariuki, S.M., Matuja, W., Akpalu, A., Kakooza-Mwesige, A., Chabi, M., Wagner, R.G., Connor, M., Chengo, E., Ngugi, A.K., Odhiambo, R., et al. (2014). Clinical features, proximate causes, and consequences of active convulsive epilepsy in Africa. Epilepsia *55*, 76–85. 10.1111/epi.12392.

3. Ba-Diop, A., Marin, B., Druet-Cabanac, M., Ngoungou, E.B., Newton, C.R., and Preux, P.M. (2014). Epidemiology, causes, and treatment of epilepsy in sub-Saharan Africa. Lancet Neurol *13*, 1029–1044. 10.1016/S1474-4422(14)70114-0.

4. Esterhuizen, A.I., Carvill, G.L., Ramesar, R.S., Kariuki, S.M., Newton, C.R., Poduri, A., and Wilmshurst, J.M. (2018). Clinical application of epilepsy genetics in Africa: Is now the time? Front Neurol *9*. 10.3389/fneur.2018.00276.

5. Adjei, P., Akpalu, A., Laryea, R., Nkromah, K., Sottie, C., Ohene, S., and Osei, A. (2013). Beliefs on epilepsy in Northern Ghana. Epilepsy and Behavior *29*, 316–321. 10.1016/j.yebeh.2013.07.034.

6. Ngugi, A.K., Bottomley, C., Kleinschmidt, I., Wagner, R.G., Kakooza-Mwesige, A., Ae-Ngibise, K., Owusu-Agyei, S., Masanja, H., Kamuyu, G., Odhiambo, R., et al. (2013). Prevalence of active convulsive epilepsy in sub-Saharan Africa and associated risk factors: Cross-sectional and case-control studies. Lancet Neurol *12*, 253–263. 10.1016/S1474-4422(13)70003-6.

7. Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S.Y., et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. The Lancet *380*, 2095–2128. 10.1016/S0140-6736(12)61728-0.

8. Meyer, A.C., Dua, T., Ma, J., Saxena, S., and Birbeck, G. (2010). Global disparities in the epilepsy treatment gap: A systematic review. Bull World Health Organ *88*, 260–266. 10.2471/BLT.09.064147.

9. Dagar, A., and Falcone, T. (2020). Psychiatric Comorbidities in Pediatric Epilepsy. Curr Psychiatry Rep *22*. 10.1007/s11920-020-01195-8.

10. Park, J.T. (2020). The epileptogenic zone in pediatric epilepsy surgery: Special issue. Seizure *77*, 1–2. 10.1016/j.seizure.2019.12.018.

11. Duchowny, M. (2020). Pediatric epilepsy surgery: Past, present, and future. Epilepsia *61*, 228–229. 10.1111/epi.16436.

12. Magiorkinis, E., Sidiropoulou, K., and Diamantis, A. (2010). Hallmarks in the history of epilepsy: Epilepsy in antiquity. Epilepsy and Behavior *17*, 103–108. 10.1016/j.yebeh.2009.10.023.

13. The International League Against Epilepsy Consortium on Complex Epilepsies (2018). Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. Nat Commun. 10.1038/s41467-018-07524-z.

14. Fisher, R.S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J.H., Elger, C.E., Engel, J., Forsgren, L., French, J.A., Glynn, M., et al. (2014). ILAE Official Report: A practical clinical definition of epilepsy. Epilepsia *55*, 475–482. 10.1111/epi.12550.

15. Scheffer, I.E., Berkovic, S., Capovilla, G., Connolly, M.B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G.W., Moshé, S.L., et al. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia *58*, 512–521. 10.1111/epi.13709.

16. Maya Kaye, A. (2021). Pediatric epilepsy and psychoeducational interventions: A review of the literature. Epilepsy and Behavior *121*. 10.1016/j.yebeh.2021.108084.

17. Fisher, R.S., Cross, J.H., French, J.A., Higurashi, N., Hirsch, E., Jansen, F.E., Lagae, L., Moshé, S.L., Peltola, J., Roulet Perez, E., et al. (2017). Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. Epilepsia *58*, 522–530. 10.1111/epi.13670.

18. Wilmshurst, J.M., Gaillard, W.D., Vinayan, K.P., Tsuchida, T.N., Plouin, P., van Bogaert, P., Carrizosa, J., Elia, M., Craiu, D., Jovic, N.J., et al. (2015). Summary of recommendations for the management of infantile seizures: Task Force Report for the ILAE Commission of Pediatrics. Epilepsia *56*, 1185–1197. 10.1111/epi.13057.

19. Specchio, N., Wirrell, E.C., Scheffer, I.E., Nabbout, R., Riney, K., Samia, P., Guerreiro, M., Gwer, S., Zuberi, S.M., Wilmshurst, J.M., et al. (2022). International League Against Epilepsy classification and definition of epilepsy syndromes with onset in childhood: Position paper by the ILAE Task Force on Nosology and Definitions. Epilepsia *63*, 1398–1442. 10.1111/epi.17241.

20. Helbig, I., and Abou Tayoun, A.N. (2016). Understanding genotypes and phenotypes in epileptic encephalopathies. Mol Syndromol *7*, 172–181. 10.1159/000448530.

21. Löscher, W., Klotz, U., Zimprich, F., and Schmidt, D. (2009). The clinical impact of pharmacogenetics on the treatment of epilepsy. Epilepsia *50*, 1–23. 10.1111/j.1528-1167.2008.01716.x.

22. Kwong, A.K.Y., Ho, A.C.C., Fung, C.W., and Wong, V.C.N. (2015). Analysis of mutations in 7 genes associated with neuronal excitability and synaptic transmission in a cohort of children with non-syndromic infantile epileptic encephalopathy. PLoS One *10*. 10.1371/journal.pone.0126446.

23. Bayat, A., Bayat, M., Rubboli, G., and Møller, R.S. (2021). Epilepsy syndromes in the first year of life and usefulness of genetic testing for precision therapy. Genes (Basel) *12*. 10.3390/genes12071051.

24. Bayat, A., Bayat, M., Rubboli, G., and Møller, R.S. (2021). Epilepsy syndromes in the first year of life and usefulness of genetic testing for precision therapy. Genes (Basel) *12*. 10.3390/genes12071051.

25. Scheffer, I.E., and Liao, J. (2020). Deciphering the concepts behind “Epileptic encephalopathy” and “Developmental and epileptic encephalopathy.” European Journal of Paediatric Neurology *24*, 11–14. 10.1016/j.ejpn.2019.12.023.

26. McTague, A., Howell, K.B., Cross, J.H., Kurian, M.A., and Scheffer, I.E. (2016). The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol *15*, 304–316. 10.1016/S1474-4422(15)00250-1.

27. Hebbar, M., and Mefford, H.C. (2020). Recent advances in epilepsy genomics and genetic testing. F1000Res *9*. 10.12688/f1000research.21366.1.

28. Symonds, J.D., and McTague, A. (2020). Epilepsy and developmental disorders: Next generation sequencing in the clinic. European Journal of Paediatric Neurology *24*, 15–23. 10.1016/j.ejpn.2019.12.008.

29. Wang, J., Lin, Z.J., Liu, L., Xu, H.Q., Shi, Y.W., Yi, Y.H., He, N., and Liao, W.P. (2017). Epilepsy-associated genes. Seizure *44*, 11–20. 10.1016/j.seizure.2016.11.030.

30. Fukata, Y., Chen, X., Chiken, S., Hirano, Y., Yamagata, A., and Inahashi, H. (2021). LGI1–ADAM22–MAGUK configures transsynaptic nanoalignment for synaptic transmission and epilepsy prevention. PNAS *118*, e2022580118. 10.1073/pnas.2022580118/-/DCSupplemental.Published.

31. Saffari, A., Schröter, J., Garbade, S.F., Alecu, J.E., Ebrahimi-Fakhari, D., Hoffmann, G.F., Kölker, S., Ries, M., and Syrbe, S. (2021). Quantitative retrospective natural history modeling of WDR45-related developmental and epileptic encephalopathy–a systematic cross-sectional analysis of 160 published cases. Autophagy. 10.1080/15548627.2021.1990671.

32. McTague, A., Howell, K.B., Cross, J.H., Kurian, M.A., and Scheffer, I.E. (2016). The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol *15*, 304–316. 10.1016/S1474-4422(15)00250-1.

33. Helbig, I., Heinzen, E.L., and Mefford, H.C. (2018). Genetic literacy series : Primer part 2 — Paradigm shifts in epilepsy genetics. 1138–1147. 10.1111/epi.14193.

34. Symonds, J.D., Zuberi, S.M., Stewart, K., McLellan, A., O'Regan, M., MacLeod, S., Jollands, A., Joss, S., Kirkpatrick, M., Brunklaus, A., et al. (2019). Incidence and phenotypes of childhood-onset genetic epilepsies: a prospective population-based national cohort. Brain *142*, 2303–2318. 10.1093/brain/awz195.

35. Heyne, H.O., Artomov, M., Battke, F., Bianchini, C., Smith, D.R., Liebmann, N., Tadigotla, V., Stanley, C.M., Lal, D., Rehm, H., et al. (2019). Targeted gene sequencing in 6994 individuals with neurodevelopmental disorder with epilepsy. Genetics in Medicine. 10.1038/s41436-019-0531-0.

36. Feng, Y.C.A., Howrigan, D.P., Abbott, L.E., Tashman, K., Cerrato, F., Singh, T., Heyne, H., Byrnes, A., Churchhouse, C., Watts, N., et al. (2019). Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. Am J Hum Genet *105*, 267–282. 10.1016/j.ajhg.2019.05.020.

37. Myers, K.A., and Scheffer, I.E. (2022). Precision Medicine Approaches for Infantile-Onset Developmental and Epileptic Encephalopathies. Annu Rev Pharmacol Toxicol, 641–662. 10.1146/annurev-pharmtox-052120.

38. Ellis, C.A., Petrovski, S., and Berkovic, S.F. (2020). Epilepsy genetics: clinical impacts and biological insights. Lancet Neurol *19*, 93–100. 10.1016/S1474-4422(19)30269-8.

39. Kaplanis, J., Samocha, K.E., Wiel, L., Zhang, Z., Arvai, K.J., Eberhardt, R.Y., Gallone, G., Lelieveld, S.H., Martin, H.C., McRae, J.F., et al. (2020). Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nature *586*, 757–762. 10.1038/s41586-020-2832-5.

40. Samanta, D. (2021). Epilepsy in Angelman syndrome: A scoping review. Brain Dev *43*, 32–44. 10.1016/j.braindev.2020.08.014.

41. Stamberger, H., Nikanorova, M., Accorsi, P., Angriman, M., Benkel-herrenbrueck, I., Capovilla, G., Erasmus, C.E., Fannemel, M., Giordano, L., Helbig, K.L., et al. (2016). STXBP1 encephalopathy: A neurodevelopmental disorder including epilepsy. Neurology *86*, 954–962.

42. Helbig, I., and Abou Tayoun, A.N. (2016). Understanding genotypes and phenotypes in epileptic encephalopathies. Mol Syndromol *7*, 172–181. 10.1159/000448530.

43. el Ghaleb, Y., Schneeberger, P.E., Fernández-Quintero, M.L., Geisler, S.M., Pelizzari, S., Polstra, A.M., van Hagen, J.M., Denecke, J., Campiglio, M., Liedl, K.R., et al. (2021). CACNA1I gain-of-function mutations differentially affect channel gating and cause neurodevelopmental disorders. Brain *144*, 2092–2106. 10.1093/brain/awab101.

44. Lewis-Smith, D., Galer, P.D., Balagura, G., Kearney, H., Ganesan, S., Cosico, M., O'Brien, M., Vaidiswaran, P., Krause, R., Ellis, C.A., et al. (2021). Modeling seizures in the Human Phenotype Ontology according to contemporary ILAE concepts makes big phenotypic data tractable. Epilepsia *62*, 1293–1305. 10.1111/epi.16908.

45. Crawford, K., Xian, J., Helbig, K.L., Galer, P.D., Parthasarathy, S., Lewis-Smith, D., Kaufman, M.C., Fitch, E., Ganesan, S., O'Brien, M., et al. (2021). Computational analysis of 10,860 phenotypic annotations in individuals with SCN2A-related disorders. Genetics in Medicine *23*, 1263–1272. 10.1038/s41436-021-01120-1.

46. Krey, I., Platzer, K., Esterhuizen, A., Berkovic, S.F., Helbig, I., Hildebrand, M.S., Lerche, H., Lowenstein, D., Møller, R.S., Poduri, A., et al. (2022). Current practice in diagnostic genetic testing of the epilepsies. Epileptic Disord *24*, 1–22. 10.1684/epd.2022.1448.

47. Musto, E., Gardella, E., and Møller, R.S. (2020). Recent advances in treatment of epilepsy-related sodium channelopathies. European Journal of Paediatric Neurology *24*, 123–128. 10.1016/j.ejpn.2019.12.009.

48. Bagnasco, I., Dassi, P., Blé, R., and Vigliano, P. (2018). A relatively mild phenotype associated with mutation of SCN8A. Seizure: European Journal of Epilepsy *56*, 47–49. 10.1016/j.seizure.2018.01.021.

49. Knowles, J.K., Helbig, I., Metcalf, C.S., Lubbers, L.S., Isom, L.L., Demarest, S., Goldberg, E.M., George, A.L., Lerche, H., Weckhuysen, S., et al. (2022). Precision medicine for genetic epilepsy on the horizon: Recent advances, present challenges, and suggestions for continued progress. Epilepsia *63*, 2461–2475. 10.1111/epi.17332.

50. Stosser, M.B., Lindy, A.S., Butler, E., Retterer, K., Piccirillo-Stosser, C.M., Richard, G., and McKnight, D.A. (2018). High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. Genetics in Medicine *20*, 403–410. 10.1038/gim.2017.114.

51. Lange, I.M. de, Koudijs, M.J., Slot, R. van, Sonsma, A.C.M., Mulder, F., Carbo, E.C., Kempen, M.J.A. van, Nijman, I.J., Ernst, R.F., Savelberg, S.M.C., et al. (2019). Assessment of parental mosaicism in SCN1A -related epilepsy by single-molecule molecular inversion probes and next-generation sequencing. J Med Genet, 75–80. 10.1136/jmedgenet-2018-105672.

52. Achkar, C.M. el, Olson, H.E., Poduri, A., and Pearl, P.L. (2015). The Genetics of the Epilepsies. 10.1007/s11910-015-0559-8.

53. Pederick, D.T., Richards, K.L., Piltz, S.G., Kumar, R., Mincheva-Tasheva, S., Mandelstam, S.A., Dale, R.C., Scheffer, I.E., Gecz, J., Petrou, S., et al. (2018). Abnormal Cell Sorting Underlies the Unique X-Linked Inheritance of PCDH19 Epilepsy. Neuron *97*, 59-66.e5. 10.1016/j.neuron.2017.12.005.

54. Helbig, I., Riggs, E.R., Barry, C.-A., Klein, K.M., Dyment, D., Thaxton, C., Sadikovic, B., Sands, T.T., Wagnon, J.L., Liaquat, K., et al. (2018). The ClinGen Epilepsy Gene Curation Expert Panel-Bridging the divide between clinical domain knowledge and formal gene curation criteria. Hum Mutat *39*, 1476–1484. 10.1002/humu.23632.

55. Esterhuizen, A.I., Mefford, H.C., Ramesar, R.S., Wang, S., Carvill, G.L., and Wilmshurst, J.M. (2018). Dravet syndrome in South African infants: Tools for an early diagnosis. Seizure *62*, 99–105. 10.1016/j.seizure.2018.09.010.

56. Du, X., Chen, Y., Zhao, Y., Luo, W., Cen, Z., and Hao, W. (2017). Dramatic response to pyridoxine in a girl with absence epilepsy with ataxia caused by a de novo CACNA1A mutation. Seizure *45*, 189–191. 10.1016/j.seizure.2016.12.020.

57. Esterhuizen, A.I., Mefford, H.C., Ramesar, R.S., Carvill, G.L., Wilmshurst, J.M., and Wang, S. (2018). Seizure : European Journal of Epilepsy Dravet syndrome in South African infants : Tools for an early diagnosis. Seizure: European Journal of Epilepsy *62*, 99–105. 10.1016/j.seizure.2018.09.010.

58. Chiron, C., and Dulac, O. (2011). The pharmacologic treatment of Dravet syndrome. Epilepsia *52*, 72–75. 10.1111/j.1528-1167.2011.03007.x.

59. Scheffer, I.E., and Nabbout, R. (2019). SCN1A-related phenotypes: Epilepsy and beyond. Epilepsia *60*, S17–S24. 10.1111/epi.16386.

60. Brunklaus, A., Brünger, T., Feng, T., Fons, C., Lehikoinen, A., Panagiotakaki, E., Vintan, M.A., Symonds, J., Andrew, J., Arzimanoglou, A., et al. (2022). The gain of function SCN1A disorder spectrum: novel epilepsy phenotypes and therapeutic implications. Brain *145*, 3816–3831. 10.1093/brain/awac210.

61. Ogiwara, I., Nakayama, T., Yamagata, T., Ohtani, H., Mazaki, E., Tsuchiya, S., Inoue, Y., and Yamakawa, K. (2012). A homozygous mutation of voltage-gated sodium channel βi gene SCN1B in a patient with Dravet syndrome. Epilepsia *53*, 200–203. 10.1111/epi.12040.

62. Ramadan, W., Patel, N., Anazi, S., Kentab, A.Y., Bashiri, F.A., Hamad, M.H., Jad, L., Salih, M.A., Alsaif, H., Hashem, M., et al. (2017). Confirming the recessive inheritance of SCN1B mutations in developmental epileptic encephalopathy. Clin Genet *92*, 327–331. 10.1111/cge.12999.

63. Wallace, R.H., Wang, D.W., Singh, R., Scheffer, I.E., George, A.L., Phillips, H.A., Saar, K., Reis, A., Johnson, E.W., Sutherland, G.R., et al. (1998). Febrile seizures and generalized epilepsy

associated with a mutation in the Na+-channel ß1 subunit gene SCN1B. Nat Genet *19*, 366–370. 10.1038/1252.

64. Steel, D., Symonds, J.D., Zuberi, S.M., and Brunklaus, A. (2017). Dravet syndrome and its mimics: Beyond SCN1A. Epilepsia *58*, 1807–1816. 10.1111/epi.13889.

65. Sisodiya, S.M. (2021). Precision medicine and therapies of the future. Epilepsia *62*, S90–S105. 10.1111/epi.16539.

66. Berg, A.T., Coryell, J., Saneto, R.P., Grinspan, Z.M., Alexander, J.J., Kekis, M., Sullivan, J.E., Wirrell, E.C., Shellhaas, R.A., Mytinger, J.R., et al. (2017). Early-life epilepsies and the emerging role of genetic testing. JAMA Pediatr *171*, 863–871. 10.1001/jamapediatrics.2017.1743.

67. Carvill, G.L., Heavin, S.B., Yendle, S.C., McMahon, J.M., O'Roak, B.J., Cook, J., Khan, A., Dorschner, M.O., Weaver, M., Calvert, S., et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet *45*, 825–830. 10.1038/ng.2646.

68. Reddy, C., and Saini, A.G. (2021). Metabolic Epilepsy. Indian J Pediatr *88*, 1025–1032. 10.1007/s12098-020-03510-w.

69. Helbig, I., Heinzen, E.L., Mefford, H.C., Berkovic, S.F., Lowenstein, D.H., Kato, M., Cross, J.H., Satishchandra, P., de Jonghe, P., Jiang, Y., et al. (2018). Genetic literacy series: Primer part 2—Paradigm shifts in epilepsy genetics. Epilepsia *59*, 1138–1147. 10.1111/epi.14193.

70. Symonds, J.D., and McTague, A. (2020). Epilepsy and developmental disorders: Next generation sequencing in the clinic. European Journal of Paediatric Neurology *24*, 15–23. 10.1016/j.ejpn.2019.12.008.

71. Liu, J., Tong, L., Song, S., Niu, Y., Li, J., Wu, X., Zhang, J., Zai, C.C., Luo, F., Wu, J., et al. (2018). Novel and de novo mutations in pediatric refractory epilepsy. Mol Brain *11*. 10.1186/s13041-018-0392-5.

72. Kothur, K., Holman, K., Farnsworth, E., Ho, G., Lorentzos, M., Troedson, C., Gupta, S., Webster, R., Procopis, P.G., Menezes, M.P., et al. (2018). Diagnostic yield of targeted massively parallel sequencing in children with epileptic encephalopathy. Seizure *59*, 132–140. 10.1016/j.seizure.2018.05.005.

73. Chambers, C., Jansen, L.A., and Dhamija, R. (2016). Review of Commercially Available Epilepsy Genetic Panels. J Genet Couns *25*, 213–217. 10.1007/s10897-015-9906-9.

74. Mercimek-Mahmutoglu, S., Patel, J., Cordeiro, D., Hewson, S., Callen, D., Donner, E.J., Hahn, C.D., Kannu, P., Kobayashi, J., Minassian, B.A., et al. (2015). Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. Epilepsia *56*, 707–716. 10.1111/epi.12954.

75. Symonds, J.D., Zuberi, S.M., Stewart, K., McLellan, A., O'Regan, M., MacLeod, S., Jollands, A., Joss, S., Kirkpatrick, M., Brunklaus, A., et al. (2019). Incidence and phenotypes of childhood-onset genetic epilepsies: A prospective population-based national cohort. Brain *142*, 2303–2318. 10.1093/brain/awz195.

76. Zeng, Q., Yang, X., Zhang, J., Liu, A., Yang, Z., Liu, X., Wu, Y., Wu, X., Wei, L., and Zhang, Y. (2018). Genetic analysis of benign familial epilepsies in the first year of life in a Chinese cohort. J Hum Genet *63*, 9–18. 10.1038/s10038-017-0359-x.

77. Jiang, Y. li, Yuan, F., Yang, Y., Sun, X. long, Song, L., and Jiang, W. (2018). CHRNA4 variant causes paroxysmal kinesigenic dyskinesia and genetic epilepsy with febrile seizures plus? Seizure *56*, 88–91. 10.1016/j.seizure.2018.02.005.

78. Allen, N.M., Conroy, J., Shahwan, A., Lynch, B., Correa, R.G., Pena, S.D.J., McCreary, D., Magalhães, T.R., Ennis, S., Lynch, S.A., et al. (2016). Unexplained early onset epileptic encephalopathy: Exome screening and phenotype expansion. Epilepsia *57*, e12–e17. 10.1111/epi.13250.

79. Fortin, O., Vincelette, C., Chénier, S., Ghais, A., Shevell, M.I., Simard-Tremblay, E., and Myers, K.A. (2020). Copy number variation in genetic epilepsy with febrile seizures plus. European Journal of Paediatric Neurology *27*, 111–115. 10.1016/j.ejpn.2020.05.005.

80. Niestroj, L.M., Perez-Palma, E., Howrigan, D.P., Zhou, Y., Cheng, F., Saarentaus, E., Nürnberg, P., Stevelink, R., Daly, M.J., Palotie, A., et al. (2020). Epilepsy subtype-specific copy number burden observed in a genome-wide study of 17458 subjects. Brain *143*, 2106–2118. 10.1093/brain/awaa171.

81. Coughlin, C.R., Tseng, L.A., Abdenur, J.E., Ashmore, C., Boemer, F., Bok, L.A., Boyer, M., Buhas, D., Clayton, P.T., Das, A., et al. (2021). Consensus guidelines for the diagnosis and management of pyridoxine-dependent epilepsy due to α-aminoadipic semialdehyde dehydrogenase deficiency. J Inherit Metab Dis *44*, 178–192. 10.1002/jimd.12332.

82. Reddy, C., and Saini, A.G. (2021). Metabolic Epilepsy. Indian J Pediatr *88*, 1025–1032. 10.1007/s12098-020-03510-w.

83. Carvill, G.L., Heavin, S.B., Yendle, S.C., McMahon, J.M., O'Roak, B.J., Cook, J., Khan, A., Dorschner, M.O., Weaver, M., Calvert, S., et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet *45*, 825–830. 10.1038/ng.2646.

84. Smith, L.A., Ullmann, J.F.P., Olson, H.E., Achkar, C.M.E., Truglio, G., Kelly, M., Rosen-Sheidley, B., and Poduri, A. (2017). A Model Program for Translational Medicine in Epilepsy Genetics. J Child Neurol *32*, 429–436. 10.1177/0883073816685654.

85. Gerbatin, R.R., Augusto, J., Morris, G., Campbell, A., Worm, J., Langa, E., Reschke, C.R., and Henshall, D.C. (2022). Investigation of MicroRNA-134 as a Target against Seizures and SUDEP in a Mouse Model of Dravet Syndrome. eNeuro *9*. 10.1523/ENEURO.0112-22.2022.

86. Vanoye, C.G., Desai, R.R., Ji, Z., Adusumilli, S., Jairam, N., Ghabra, N., Joshi, N., Fitch, E., Helbig, K.L., McKnight, D., et al. (2022). High-throughput evaluation of epilepsy-associated KCNQ2 variants reveals functional and pharmacological heterogeneity. JCI Insight. 10.1172/jci.insight.156314.

87. Helbig, I., and Ellis, C.A. (2020). Personalized medicine in genetic epilepsies – possibilities, challenges, and new frontiers. Neuropharmacology *172*. 10.1016/j.neuropharm.2020.107970.

88. Pohlmann-Eden, B., and Weaver, D.F. (2013). The puzzle(s) of pharmacoresistant epilepsy. Epilepsia *54 Suppl 2*, 1–4. 10.1111/epi.12174.

89. Kwan, P., Arzimanoglou, A., Berg, A.T., Brodie, M.J., Hauser, W.A., Mathern, G., Moshé, S.L., Perucca, E., Wiebe, S., and French, J. (2010). Definition of drug resistant epilepsy: Consensus

proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. Epilepsia *51*, 1069–1077. 10.1111/j.1528-1167.2009.02397.x.

90. Lopez-garcia, M.A., Feria-romero, I.A., Serrano, H.F.-, Escalante, D., and Suarez, S.O.- (2014). Genetic polymorphisms associated with antiepileptic metabolism. Frontiers in Bioscience, 377–386.

91. Balestrini, S., and Sisodiya, S.M. (2018). Pharmacogenomics in epilepsy. Neurosci Lett *667*, 27–39. 10.1016/j.neulet.2017.01.014.

92. Verrotti, A., Grasso, E.A., Cacciatore, M., Matricardi, S., and Striano, P. (2021). Potential role of brivaracetam in pediatric epilepsy. Acta Neurol Scand *143*, 19–26. 10.1111/ane.13347.

93. Fricke-Galindo, I., Jung-Cook, H., LLerena, A., and López-López, M. (2018). Pharmacogenetics of adverse reactions to antiepileptic drugs. Neurología (English Edition) *33*, 165–176. 10.1016/j.nrleng.2015.03.021.

94. Hockings, J.K., Pasternak, A.L., Erwin, A.L., Mason, N.T., Eng, C., and Hicks, J.K. (2020). Pharmacogenomics: An evolving clinical tool for precision medicine. Cleve Clin J Med *87*, 91–99. 10.3949/ccjm.87a.19073.

95. Rajman, I., Knapp, L., Morgan, T., and Masimirembwa, C. (2017). African Genetic Diversity: Implications for Cytochrome P450-mediated Drug Metabolism and Drug Development. EBioMedicine *17*, 67–74. 10.1016/j.ebiom.2017.02.017.

96. Esterhuizen, A.I., Tiffin, N., Riordan, G., Wessels, M., Burman, R.J., Aziz, M.C., Calhoun, J.D., Gunti, J., Amiri, E.E., Ramamurthy, A., et al. (2022). Precision medicine for developmental and epileptic encephalopathies in Africa—strategies for a resource-limited setting. Genetics in Medicine. 10.1016/j.gim.2022.11.002.

97. Wang, J., Lin, Z.J., Liu, L., Xu, H.Q., Shi, Y.W., Yi, Y.H., He, N., and Liao, W.P. (2017). Epilepsy-associated genes. Seizure *44*, 11–20. 10.1016/j.seizure.2016.11.030.

98. Perucca, P., and Perucca, E. (2019). Identifying mutations in epilepsy genes: Impact on treatment selection. Epilepsy Res *152*, 18–30. 10.1016/j.eplepsyres.2019.03.001.

99. Verma, V., Gupta, A., and Chaudhary, V.K. (2020). Emulsion PCR made easy. Biotechniques *69*, 65–69. 10.2144/BTN-2019-0161.

100. Breda Genetics (2021). Coverage in next-generation sequencing.

101. Homer, N., Merriman, B., and Nelson, S.F. (2009). Local alignment of two-base encoded DNA sequence. BMC Bioinformatics *10*. 10.1186/1471-2105-10-175.

102. Homer, N., Merriman, B., and Nelson, S.F. (2009). BFAST: An alignment tool for large scale genome resequencing. PLoS One *4*. 10.1371/journal.pone.0007767.

103. Mani, A. (2017). Pathogenicity of de Novo Rare Variants: Challenges and Opportunities. Circ Cardiovasc Genet *10*. 10.1161/CIRCGENETICS.117.002013.

104. Sims, D., Sudbery, I., Ilott, N.E., Heger, A., and Ponting, C.P. (2014). Sequencing depth and coverage: Key considerations in genomic analyses. Nat Rev Genet *15*, 121–132. 10.1038/nrg3642.

105. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine *17*, 405–424. 10.1038/gim.2015.30.

106. Gudiseva, H. v., Hansen, M., Gutierrez, L., Collins, D.W., He, J., Verkuil, L.D., Danford, I.D., Sagaser, A., Bowman, A.S., Salowe, R., et al. (2016). Saliva DNA quality and genotyping efficiency in a predominantly elderly population. BMC Med Genomics *9*. 10.1186/s12920-016-0172-y.

107. Thermo Fisher Scientific T042-TECHNICAL BULLETIN NanoDrop Spectrophotometers.

108. Melamed, D., Nov, Y., Malik, A., Yakass, M.B., Bolotin, E., Shemer, R., Hiadzi, E.K., Skorecki, K.L., and Livnat, A. De novo mutation rates at the single-mutation resolution in a human HBB gene-region associated with adaptation and genetic disease.

109. Coughlin, C.R., Tseng, L.A., Abdenur, J.E., Ashmore, C., Boemer, F., Bok, L.A., Boyer, M., Buhas, D., Clayton, P.T., Das, A., et al. (2021). Consensus guidelines for the diagnosis and management of pyridoxine-dependent epilepsy due to α-aminoadipic semialdehyde dehydrogenase deficiency. J Inherit Metab Dis *44*, 178–192. 10.1002/jimd.12332.

110. Oegema, R., Barakat, T.S., Wilke, M., Stouffs, K., Amrom, D., Aronica, E., Bahi-Buisson, N., Conti, V., Fry, A.E., Geis, T., et al. (2020). International consensus recommendations on the diagnostic work-up for malformations of cortical development. Nat Rev Neurol *16*, 618–635. 10.1038/s41582-020-0395-6.

111. Cohen, M.M. (2013). The AKT genes and their roles in various disorders. Am J Med Genet A *161*, 2931–2937. 10.1002/ajmg.a.36101.

112. Franz, D.N., and Weiss, B.D. (2012). Molecular therapies for tuberous sclerosis and neurofibromatosis. Curr Neurol Neurosci Rep *12*, 294–301. 10.1007/s11910-012-0269-4.

113. Scheffer, I.E., Heron, S.E., Regan, B.M., Mandelstam, S., Crompton, D.E., Hodgson, B.L., Sci, B., Licchetta, L., Provini, F., Bisulli, F., et al. (2014). Mutations in Mammalian Target of Rapamycin Regulator DEPDC5 Cause Focal Epilepsy with Brain Malformations Patients and Methods Phenotyping. American Neurological Association, 782–787. 10.1002/ana.

114. Shirley, M.D., Tang, H., Gallione, C.J., Baugher, J.D., Frelin, L.P., Cohen, B., North, P.E., Marchuk, D.A., Comi, A.M., and Pevsner, J. (2013). Sturge–Weber Syndrome and Port-Wine Stains Caused by Somatic Mutation in GNAQ . New England Journal of Medicine *368*, 1971–1979. 10.1056/nejmoa1213507.

115. Shestak, A.G., Bukaeva, A.A., Saber, S., and Zaklyazminskaya, E. v. (2021). Allelic Dropout Is a Common Phenomenon That Reduces the Diagnostic Yield of PCR-Based Sequencing of Targeted Gene Panels. Front Genet *12*. 10.3389/fgene.2021.620337.

116. Benmakhlouf, Y., Touraine, R., Harzallah, I., Zian, Z., ben Makhlouf, K., Barakat, A., Ghailani Nourouti, N., and Bennani Mechita, M. (2021). Screening of the duplication 24 pb of ARX gene in Moroccan patients with X-linked Intellectual Disability. BMC Res Notes *14*. 10.1186/s13104-021-05526-7.

117. Shoubridge, C., Jackson, M., Grinton, B., Berkovic, S.F., Scheffer, I.E., Huskins, S., Thomas, A., and Ware, T. (2019). Splice variant in ARX leading to loss of C-terminal region in a boy with intellectual disability and infantile onset developmental and epileptic encephalopathy. Am J Med Genet A *179*, 1483–1490. 10.1002/ajmg.a.61216.

118. Wong, L.C., Singh, S., Wang, H.P., Hsu, C.J., Hu, S.C., and Lee, W.T. (2019). FOXG1-related syndrome: From clinical to molecular genetics and pathogenic mechanisms. Int J Mol Sci *20*. 10.3390/ijms20174176.

119. Li, D., Yuan, H., Ortiz-Gonzalez, X.R., Marsh, E.D., Tian, L., McCormick, E.M., Kosobucki, G.J., Chen, W., Schulien, A.J., Chiavacci, R., et al. (2016). GRIN2D Recurrent De Novo Dominant Mutation Causes a Severe Epileptic Encephalopathy Treatable with NMDA Receptor Channel Blockers. Am J Hum Genet *99*, 802–816. 10.1016/j.ajhg.2016.07.013.

120. Johannesen, K.M., Gardella, E., Gjerulfsen, C.E., Bayat, A., Rouhl, R.P.W., Reijnders, M., Whalen, S., Keren, B., Buratti, J., Courtin, T., et al. (2021). PURA-Related Developmental and Epileptic Encephalopathy Phenotypic and Genotypic Spectrum. Neurol Genet *7*. 10.1212/NXG.0000000000000613.

121. Clinical and Laboratory Standards Institute Quality Management for Molecular Genetic Testing; Approved Guideline.

122. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine *17*, 405–424. 10.1038/gim.2015.30.

123. Chambers, C., Jansen, L.A., and Dhamija, R. (2016). Review of Commercially Available Epilepsy Genetic Panels. J Genet Couns *25*, 213–217. 10.1007/s10897-015-9906-9.

124. Riggs, E.R., Andersen, E.F., Cherry, A.M., Kantarci, S., Kearney, H., Patel, A., Raca, G., Ritter, D.I., South, S.T., Thorland, E.C., et al. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genetics in Medicine *22*, 245–257. 10.1038/s41436-019-0686-8.

125. Mullen, S.A., Carvill, G.L., Bellows, S., Bayly, M.A., Berkovic, S.F., Dibbens, L.M., Scheffer, I.E., and Mefford, H.C. (2013). Copy number variants are frequent in genetic generalized epilepsy with intellectual disability.

126. Kane, M. (2021). CYP2D6 Overview: Allele and Phenotype Frequencies.

127. Rogawski, M.A., and Löscher, W. (2004). The neurobiology of antiepileptic drugs. Nat Rev Neurosci *5*, 553–564. 10.1038/nrn1430.

Appendix A – Patient recruitment information pamphlet (English version)

PRECISION MANAGEMENT OF EPILEPSY IN SOUTH AFRICAN CHILDREN (PME STUDY)

PATIENT INFORMATION SHEET

What you should know about this research study:

We give you this information so that you may read about the purpose, nature, risks, and benefits of this research study.

- Routine care is based upon the best available treatment and is provided with the goal of helping the individual patient.
- The main goal of research studies is to gain knowledge that may help future patients.
- We cannot promise that this research will benefit you or your child.
- Your participation is voluntary.
- You have the right to refuse to take part, or agree to take part now and change your mind later.
- Whatever you decide, it will not affect your child's regular care.

Please review this information form carefully. Ask any questions you have before you make a decision.

Purpose of this study:

You are being asked to allow your child to participate in a research study of new ways to try and understand more about difficult to control epilepsy in children. This includes closer monitoring and understanding some of the inherited contributions to the cause of epilepsy and to the body's handling of medication.

The purpose of this study is to establish if these new approaches are feasible in South Africa and show signs of being helpful in the future to improve the treatment of children with difficult to control epilepsy in our country.

Your child was selected as a possible participant in this study because he or she has complicated epilepsy which requires more than one antiepileptic drug to help in treatment.

Study procedures:

The project can be considered in two parts.

Firstly, we are looking at the feasibility of monitoring day to day events more closely using modern technology. This will include a cell phone app which will allow for close recording of events in your child's life as well as have other features such as

medication reminders. In addition, for children that will be suited to wearing a device similar to a watch, this will be used to monitor more direct information about your child.

Secondly, specific genetic tests related to epilepsy will be done on your child's DNA sample. DNA is the substance our genes and inherited material is made up of. This provides an instruction "code" that directs how our body is built and how it functions. Your child's DNA will be tested for small changes known to have an effect on how a person handles and responds to certain medication. This area of study is called pharmacogenomics and is thought to be one of the reasons why some people may not improve on a specific medication or why they may get more side effects. These changes vary between populations and very few studies have has been done in Africa so this will be valuable information.

In addition, if the cause of your child's epilepsy is unknown and thought to possibly be due to a change in a gene that has been shown to cause epilepsy, your child's DNA will also be tested for changes in these specific epilepsy genes. Knowing if there is a specific genetic cause for epilepsy may help doctors to decide on the best treatment as well as give parents more information. There is no testing available in South Africa at the moment for these genetic changes and part of this study is to develop such testing. If your child is to have epilepsy gene testing, you will also be asked for a sample in case we need to compare particular findings in your child's DNA with his or her parents in order to help decide on the significance of such genetic changes.

Although this study will not provide all the answers, we expect it will give us a good indication on how we can use these new ways of thinking about a condition like epilepsy in a South African setting to improve the way children with epilepsy can be treated in the future.

What is required from you and your child?

Your child will continue to receive all routine care for their epilepsy including medication and clinic visits. Participation in the study will not in any way change this unless information comes to light as a result of the study that would be directly beneficial to your child's care.

Home monitoring:

You will be assisted to download a cell phone app and shown how to enter information into it about your child and relevant daily events such as seizures. This can be expected to take not more than 5 minutes a day but may need a few entries to be made in a day. In addition your child will be given a home monitoring device that is similar to a watch worn on the wrist if it is thought that your child would be able to tolerate this. The device would have to be worn at all times except during bathing or swimming and while recharging the battery. It would gather information on what is happening with your child on an ongoing basis. These devices have been used before in studies in a number of countries involving many children with different conditions and have been well tolerated and not found to have any negative effects. The information collected via this monitoring will be stored on a secure database which only the study investigators will have access to.

No information that directly identifies you or your child such as names, address or date of birth will be stored on the phone app. Only the research team and your Doctor will have access to the data collected.

Samples for DNA testing

Saliva sample:

Samples for DNA extraction will be taken from your child by saliva sampling if at all possible. In older children this is simply spitting into a container but in younger children, it may mean collecting the saliva by putting small soft sponges into his or her mouth.

Blood sample:

If it is not possible to get a suitable saliva sample, we will ask to take 2-4ml (less than a teaspoonful) of blood from your child to obtain DNA. For all children, the DNA will be tested for small changes which contribute to the way in which medications are handles in the body (pharmacogenomics variants). If your child is thought to possibly have a genetic cause of their epilepsy, DNA (from the same sample) will also be tested for possible disease-causing changes in genes known to be associated with epilepsy.

After the study, extracted DNA will be stored in the DNA Registry of the Division of Human Genetics (HREC REF 217/2010) under a specific DNA code, for an unspecified period of time if you consent to such storage.

Will you be informed of the results?

No, individual results will not be provided as this is a research study to gather information rather than a test for a specific individual. Having said that, all results will be looked at by the study doctors and if there is a result that they believe is important for your child's medical care, that result will be shared with your child's doctor. We will provide feedback to you on the overall outcomes of the study through a summary report.

Voluntarily participation & confidentiality

- Participation is completely voluntary for both your child and for you.
- You have the right to refuse to participate in the study.
- A decision to refuse will not affect your child's treatment now or in the future.
- You are free to withdraw consent at any time.
- All information given will remain strictly confidential. All information and each specimen will be stored in a secure location and assigned an anonymous code known only to the investigators working on the project.
- You may request to have your child's DNA sample removed from the Repository and discarded at any time during or after completion of the study.
- This study is being conducted according to the principles of the Declaration of Helsinki (Brazil 2013), which looks after the interest of the participants.

What if something goes wrong?

If you have any problems with the device or phone app you will be asked to contact Aparito (the tech providers) on email via support@aparito.com. The research nurse or Dr Fieggen can assist you with this if necessary. We plan to disclose any relevant safety information to the Data and Safety Monitoring Committee (DSMC) as they have role to evaluate whether the study should be terminated to ensure the safety of the participants.

It is considered very unlikely for anything serious to go wrong in this study but the University of Cape Town (UCT) undertakes to provide immediate medical care in the event of your child suffering any significant deterioration in health or well-being from any unexpected sensitivity or toxicity, that is caused by his or her participation in the study. UCT has appropriate insurance and guidelines to that to manage any trial-related injury in keeping with the guidelines outlined by the Association of the British Pharmaceutical Industry, ABPI 1991. UCT reserves the right not to provide compensation if, and to the extent that, your injury came about because you chose not to follow the instructions that you were given while you were taking part in the study.

Your right in law to claim compensation for injury where you prove negligence is not affected. Copies of these guidelines are available on request.

Benefits and/or compensation

We cannot and do not guarantee or promise that you or your family will receive any benefits from this study. The research is however conducted with a view to improving care for children with epilepsy. Participation may therefore have a long-term benefit in helping to improve the diagnosis and treatment of epilepsy in South African children.

You will not be compensated or paid for taking part in the study but there will not be any additional costs to you for participation and any costs incurred as direct result of the study will be covered by the study funding.

Study discontinuation criteria

You may voluntarily withdraw your child from the study at any time. Your child may also be withdrawn from the study at the discretion of the investigators at any time. This can be due to one of the following reasons:

1. Any Serious Adverse Event that is considered possibly related to study intervention
2. Major protocol violations
3. Any other situation where the investigator believes that continuation of study is contrary to the best interest of the subject.

Future research

With your consent, your child's DNA (and yours if relevant) will be stored in case further genetic studies become available in the research study area of epilepsy. Although future research that uses your child's samples may possibly lead to the development of new products, no financial compensation will be provided to you should this occur.

Who to contact if you have any questions?

 082 396 0105  pmestudy@gmail.com

You can also contact Dr Fieggen at 021 4066298 or 0832854707 or at karen.fieggen@uct.ac.za.

Participants may contact the Human Research Ethics Committee at UCT with any questions or concerns about their rights or welfare as research participants: tel: +27 21 406 6338; fax: +27 21 406 6411; Email: nosi.tsama@uct.ac.za or shuretta.thomas@uct.ac.za

Appendix B – Manufacturer's instructions for DNA isolation from saliva collected in Oragene kits

DNA ISOLATION FROM SALIVA

Oragene Kit

1. Invert Oragene saliva tube to mix and incubate at 50°C for an hour
2. Transfer saliva into a 15 ml Greiner tube
3. Add 1/25 volume (i.e. 4 ml sample= 160 µl of PrepIt) of the Oragene DNA purifier and vortex thoroughly
4. Incubate at -20°C for 10 minutes and centrifuge at 2400 rpm for 10 minutes
5. Transfer the supernatant into a clean 15 ml tube, add 2X Volume of Absolute Ethanol and invert to mix
6. Centrifuge tubes for 10 minutes at 2400 rpm, decant the supernatant
7. Re-suspend pellet in ice cold 70% ethanol and invert to mix
8. Centrifuge at 2400 rpm for 10 minutes and decant supernatant
9. Allow the pellet to air dry at least for 2 hours
10. Reconstitute pellet in 200 µl 1X TE buffer for at least 2 days.

Appendix C – Thermo Fisher BigDye® Direct Cycle Sequencing Kit Protocol (PN 4458040C), pg. 7-17

BigDye® Direct Cycle Sequencing Kit

Product information

Purpose of the product

Use the BigDye® Direct Cycle Sequencing Kit to perform PCR, post-PCR clean-up, and cycle sequencing. The BigDye® Direct PCR Master Mix and the BigDye® Direct Sequencing Master Mix included in the kit are optimized specifically to work together. The BigDye® Direct Sequencing Master Mix contains BigDye® terminators. The sequencing master mix also allows you to combine post-PCR clean-up with cycle sequencing in the same reaction plate. The elimination of the post-PCR clean-up step simplifies and streamlines the sequencing workflow. The universal sequencing primers included in the kit enable you to perform sequencing runs with POP-7™ polymer and obtain high quality 5' data. With POP-7™ polymer, the electrophoresis times are shorter compared to runs using POP-6™ polymer.

Kit contents and storage

Three kit sizes are available.

Kit	Part no.
BigDye® Direct Cycle Sequencing Kit, 24 reactions	4458689
BigDye® Direct Cycle Sequencing Kit, 100 reactions	4458687
BigDye® Direct Cycle Sequencing Kit, 1000 reactions	4458688

Upon receipt, store the tubes at the recommended temperatures.

Component	Storage
BigDye® Direct PCR Master Mix	–20°C until first use, then 4°C
BigDye® Direct Sequencing Master Mix†	–15°C to –25°C
BigDye® Direct M13 Fwd Primer	–15°C to –25°C
BigDye® Direct M13 Rev Primer	–15°C to –25°C
Control DNA CEPH 1347-02, 20 µL, 50 ng/µL‡	–15°C to –25°C

† For optimal performance, freeze and thaw the BigDye® Direct Sequencing Master Mix no more than 5 times.

‡ The control DNA is provided at 50 ng/µL. Dilute to 4 ng/µL before use.

Materials and equipment required but not included

For capillary electrophoresis and sequencing analysis materials and equipment, see page 23.

Equipment

Item	Source
Thermal cycler: • Veriti® 96-Well Fast Thermal Cycler • Veriti® 96-Well Thermal Cycler • Aluminum 96-Well GeneAmp® PCR System 9700 • Gold-Plated 96-Well GeneAmp® PCR System 9700 • Silver 96-Well GeneAmp® PCR System 9700	Life Technologies • PN 4375305 • PN 4375786 • PN 4314879 • PN 4314878 • PN N8050001
Pipettors: 1- to 20-µL range, 20- to 200-µL range, 100- to 1000-µL range	Major laboratory supplier (MLS)
Multichannel pipettor	MLS
Plate vortexer	MLS
Tabletop centrifuge with microplate adapters	MLS
Vacuum centrifuge (optional)	MLS

Reagents and plastics

For the SDS of any chemical not distributed by Life Technologies Corporation, contact the chemical manufacturer. Before handling any chemicals, refer to the SDS provided by the manufacturer, and observe all relevant precautions.

Table 1 Reagents and plastics for PCR amplification and cycle sequencing

Item	Source
MicroAmp® 96-well optical reaction plates: • MicroAmp® Optical 96-Well Reaction Plate, 10 plates • MicroAmp® Optical 96-Well Reaction Plate with Barcode, 20 plates • MicroAmp® Optical 96-Well Reaction Plate with Barcode and Optical Adhesive Films, 100 plates with covers • MicroAmp® Optical 96-Well Reaction Plate with Barcode and Optical Caps, 20 plates with caps • MicroAmp® Fast Optical 96-Well Reaction Plate with Barcode, 0.1 mL, 20 plates • MicroAmp® Fast Optical 96-Well Reaction Plate with Barcode, 0.1 mL, 200 plates	Applied Biosystems • PN N8010560 • PN 4306737 • PN 4314320 • PN 403012 • PN 4346906 • PN 4366932
MicroAmp® caps strips: • MicroAmp® 8-Cap Strip, Assorted Colors, 300 strips • MicroAmp® 8-Cap Strip, 1500 strips • MicroAmp® 8-Cap Strip, 300 strips	Applied Biosystems • PN N8010835 • PN N8011535 • PN N8010535

Item	Source
MicroAmp® tubes and tube strips: • MicroAmp® Reaction Tube with Cap, 0.2 mL, 1000 tubes • MicroAmp® 8-Tube Strip, 0.2 mL, 125 strips • MicroAmp® Fast 8-Tube Strip, 0.1 mL, 125 strips • MicroAmp® Fast Reaction Tube with Cap, 0.1 mL, 1000 tubes	Applied Biosystems • PN N8010540 • PN N8010580 • PN 4358293 • PN 4358297
MicroAmp® Clear Adhesive Film, 200 covers	Applied Biosystems PN 4306311
MicroAmp® Optical Adhesive Film, 25 covers	Applied Biosystems PN 4360954
MicroAmp® Optical Film Compression Pad, 5 each	Applied Biosystems PN 4312639
Deionized water	MLS

Table 2 Reagents and plastics for purification of sequencing products

Item	Source
For the BigDye XTerminator® Purification Kit	
BigDye XTerminator® Purification Kit: • 2 ml (~100 20-µL reactions) • 20 ml (~1,000 20-µL reactions) • 50 ml (~2,500 20-µL reactions) • 800 ml	Applied Biosystems • PN 4376486 • PN 4376487 • PN 4376484 • PN 4376485
Heat Seal Film for Sequencing and Fragment Analysis Sample Plates, 1 package **Note:** Use for direct injection without a septa mat on the 3730/3730*xl* instrument.	Applied Biosystems PN 4337570
For spin column or spin plate purification	
2.2% Sodium dodecyl sulfate (SDS)	MLS
Centri-Sep™ Columns: • 100 columns • 32 columns	Applied Biosystems • PN 401762 • PN 401763
Centri-Sep™ 96-Well Plates: • 50 plates • 2 plates	Applied Biosystems • PN 4367821 • PN 4367819

Workflow

Use the BigDye Direct Cycle Sequencing Kit

Perform PCR amplification (page 10)

↓

Perform cycle sequencing (page 12)

↓

Purify the sequencing products (page 13)

↓

Perform capillary electrophoresis (page 16)

Perform PCR amplification

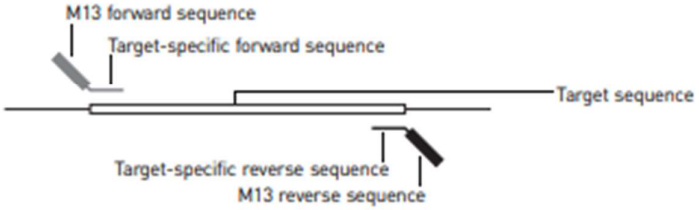
Guidelines for genomic DNA template quality

For optimal results, use high-quality gDNA with no degradation or contaminants:

- **A_{260}/A_{280} ratio**: Use gDNA with an A_{260}/A_{280} ratio between 1.7 and 1.9. A ratio outside of this range indicates contamination by protein or by organic chemicals, such as salt or phenol.
- **0.4% agarose gel**: Use gDNA that appears as a single, high-molecular-weight band (>10 kb). Smearing and low-molecular-weight bands indicate DNA degradation or RNA contamination.

PCR primer design requirements

To use the BigDye® Direct Kit, your PCR primers must include the M13 universal primer sequences as illustrated below:



- **M13 forward primer sequence**: 5′ TGTAAAACGACGGCCAGT 3′
- **M13 reverse primer sequence**: 5′ CAGGAAACAGCTATGACC 3′

Prepare and run the PCR reactions

1. For each forward or reverse reaction, add the components to an appropriate reaction plate:

Component	Volume
Genomic DNA (4 ng/µL)	1.0 µL
M13-tailed PCR primer mix (0.8 µM each primer)	1.5 µL
BigDye® Direct PCR Master Mix	5.0 µL
Deionized water	2.5 µL
Total volume for each reaction	**10.0 µL**

Note: The PCR primers must include the M13 forward and reverse sequences (page 10).

2. Pipet up and down to mix well, seal the plate with adhesive film or caps, then spin the plate briefly.

3. Run the reactions in a thermal cycler:

Stage	Veriti® thermal cyclers		9700 thermal cycler	
	Temp	Time	Temp	Time
Hold	95°C	10 min	96°C	5 min
Cycle (35 cycles)	96°C	3 sec	94°C	30 sec
	62°C	15 sec	62°C	45 sec
	68°C	30 sec	68°C	45 sec
Hold	72°C	2 min	72°C	2 min
Hold	4°C	∞	4°C	∞

4. (Optional) For replicates or controls, assess the quality and quantity of PCR products by running them on an agarose gel.

STOPPING POINT. (Optional) Store the amplified DNA at 4°C overnight or at –15°C or –25°C for long-term storage.

Perform cycle sequencing

Quantity of PCR product to use

The minimum quantity of PCR product to use for sequencing is 20 ng, as checked by running on an agarose gel.

Prepare and run the cycle sequencing reactions

IMPORTANT! You need to use the BigDye® Direct M13 forward or reverse primers in your BigDye® Direct cycle sequencing reactions.

1. Prepare a forward or reverse sequencing reaction mix in a tube on ice:

Components	Volume for each reaction
BigDye® Direct Sequencing Master Mix	2.0 µL
One sequencing primer: • BigDye® Direct M13 Fwd Primer or • BigDye® Direct M13 Rev Primer	1.0 µL
Total volume for each reaction	**3.0 µL**

2. For each sequencing reaction, add 3 µL of the sequencing reaction mix to the appropriate well in the respective forward or reverse reaction plate.
3. Seal the reaction plate with adhesive film or caps, then spin the plate briefly.
4. Run the reactions in a thermal cycler:

Stage	Veriti® thermal cyclers		9700 thermal cycler	
	Temp	Time	Temp	Time
Hold	37°C	15 min	37°C	15 min
Hold	80°C	2 min	80°C	2 min
Hold	96°C	1 min	96°C	1 min
Cycle (25 cycles)	96°C	10 sec	96°C	10 sec
	50°C	5 sec	50°C	5 sec
	60°C	75 sec	60°C	4 min
Hold	4°C	∞	4°C	∞

5. After the cycle sequencing reactions are complete, spin the plate briefly.

STOPPING POINT. (Optional) Store the reaction plate at 4°C overnight or at –15°C or –25°C for long-term storage.

Purify the sequencing products

For optimal results, completely remove unincorporated dye terminators before performing capillary electrophoresis. Excess dye terminators in sequencing products can obscure data in the early part of the sequence and interfere with basecalling.

The purification methods described here are optimized for use with the BigDye® Direct Cycle Sequencing Kit at the specified sequencing volumes and are not recommended for other BigDye® products.

Select the purification method

Select the method that works best for your application:

- BigDye XTerminator® Purification Kit (page 13)
- Spin columns (page 14)
- Spin plates (page 15)

Purify sequencing products using the BigDye XTerminator® Purification Kit

Note: This procedure is specifically optimized for use with the BigDye® Direct Cycle Sequencing Kit.

1. Spin the reaction plate at 100 *x g* for 1 minute, then remove the seal.
2. Prepare a premix with SAM™ Solution and XTerminator® Solution in an appropriately sized tube:

Component	Volume for 1 well	Volume for 96 wells
SAM™ Solution	45 µL	4752 µL
XTerminator® Solution	10 µL	1056 µL
Total volume	55 µL	5808 µL

   a. Add the SAM™ Solution to the tube using a conventional pipette tip.

      Note: Make sure there are no particulates in the SAM Solution before pipetting. If there are particulates, heat the SAM Solution to 37°C and mix to resuspend. Cool to room temperature before using.

   b. Vortex the XTerminator® Solution bulk container at maximum speed for at least 10 seconds, until the solution is homogeneous.

   c. Using a wide-bore pipette tip, aspirate the XTerminator® Solution.

      IMPORTANT! Avoid pipetting from the top of the liquid.

   d. Mix the reagents until homogeneous.

3. Add 55 µL of SAM™ Solution/XTerminator® Solution premix to each well.
4. Seal the plate using one of the following methods, then verify that each well is sealed:
   - MicroAmp® Clear Adhesive Films

     or
   - A heat seal at 160°C for 1.5 seconds

   Note: For direct injections without a septa mat on the 3730/3730*xl* instrument, only the Heat Seal Film for Sequencing and Fragment Analysis Sample Plates are supported.

5. Vortex the reaction plate for 20 minutes, using the following conditions:

Vortexer	Speed
Digital Vortex-Genie® 2	1800 rpm
IKA MS3 Digital	2000 rpm†
IKA Vortex 3	Setting 5‡
Taitec MicroMixer E-36	Maximum
Union Scientific Vertical Shaker	Setting 100§

† Set the vortexer to Mode B.
‡ Use the maximum setting without allowing the vortexer to move across the bench.
§ Add more plates, if necessary, to meet mass requirements.

6. In a swinging-bucket centrifuge, spin the plate at 1000 × *g* for 2 minutes.

STOPPING POINT. If you plan to store the plate before proceeding with capillary electrophoresis, store the sample plates sealed with heat seal film or adhesive film for up to 48 hours at room temperature (20 to 25°C) or up to 10 days at 4°C or -20°C.

Purify sequencing products with Centri-SEP™ spin columns

1. Treat the samples with 2.2% sodium dodecyl sulfate (SDS):
   a. Prepare 2.2% SDS in deionized water. This SDS solution is stable at room temperature.
   b. Add 7 µL of water and 2 µL of 2.2% SDS to each 13 µL of completed cycle sequencing reaction to bring the final SDS concentration to 0.2%.
   c. Seal the plate, then mix thoroughly.
   d. Incubate the reaction plate in a thermal cycler:

Temp	Time
98°C	5 min
25°C	10 min

   e. Spin the plate briefly.
2. Hydrate the spin columns:
   a. Gently tap the column to cause the gel material to settle to the bottom of the column.
   b. Remove the upper end cap and add 0.8 mL of deionized water.
   c. Replace the upper end cap and vortex or invert the column a few times to mix the water and gel material.
   d. Allow the gel to hydrate at room temperature for at least 2 hours.

Note: You can store hydrated columns for a few days at 2–6°C. Longer storage in water is not recommended. Allow columns stored at 2–6°C to warm to room temperature before use.

3. Assemble the spin columns:
   a. Invert or tap the column and allow the gel to settle to remove air bubbles.
   b. Remove the upper end cap first, then remove the bottom cap.
   c. Allow the column to drain completely by gravity.

      Note: If flow does not begin immediately, apply gentle pressure to the column with a pipette bulb.

   d. Insert the column into the wash tube provided.
   e. Spin the column in a microcentrifuge at 730 *x g* for 2 minutes to remove the interstitial fluid.
   f. Remove the column from the wash tube, then insert it into a sample collection tube.
4. Purify the sequencing products using the spin columns:
   a. Carefully transfer the sequencing reaction treated with SDS from its tube to the center of the gel material.
   b. Spin the column in a microcentrifuge at 730 *x g* for 2 minutes to elute the sample into the sample collection tube.

      Note: If you are using a centrifuge with a fixed-angle rotor, place the column in the same orientation as it was in for the first spin. This is important because the surface of the gel will be at an angle in the column after the first spin.

   c. Discard the column.
5. Dry the sample in a vacuum centrifuge for 10–15 minutes without heat, or until dry.

   Note: Do not overdry.

Purify sequencing products with Centri-SEP™ spin plates

For large-scale procedures, you can use Centri-SEP™ 96-well spin plates.

1. Treat the samples with 2.2% sodium dodecyl sulfate (SDS):
   a. Prepare 2.2% SDS in deionized water. This SDS solution is stable at room temperature.
   b. Add 7 µL of water and 2 µL of 2.2% SDS to each 13 µL of completed cycle sequencing reaction to bring the final SDS concentration to 0.2%.
   c. Seal the plate, then mix thoroughly.
   d. Heat the reaction plate, then allow the plate to cool to ambient temperature in the thermal cycler:

Temp	Time
98°C	5 min
25°C	10 min

   e. Spin the plate briefly.
2. Prepare the spin plate and perform the purification according to the manufacturer's instructions.

Perform capillary electrophoresis

Refer to your instrument user guide for instructions on setting up and performing the capillary electrophoresis run.

- Use Dye Set Z and the Sequencing Install Standard, BigDye® Terminator v3.1 Kit to create the BigDye® Direct spectral calibration information to apply to the data.

  Note: Current Dye Set Z spectra information can be applied to BigDye® Direct data without rerunning spectral calibration.

- Use the BigDye® mobility and calibration files for optimal basecalling with the BigDye® Direct Cycle Sequencing Kit.

  Note: For instructions on installing the mobility and calibration files, see Appendix B, "Installing the BigDye® Direct Cycle Sequencing Kit Software v1.1" on page 25.

Capillary electrophoresis run conditions

Table 3 Run conditions for 3500/3500*xl* DNA Analyzers

Polymer	Array	Run module	Mobility file
POP-7™ polymer	50 cm	StdSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	FastSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	RapidSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	ShortReadSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_StdSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_FastSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_RapidSeq50_POP7	KB_3500_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_ShortReadSeq50_POP7	KB_3500_POP7_BDTv3direct.mob

Table 4 Run conditions for 3130/3130*xl* Genetic Analyzers

Polymer	Array	Run module	Mobility file
POP-7™ polymer	36 cm	RapidSeq36_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	UltraSeq36_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	StdSeq50_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	FastSeq50_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	BDX_RapidSeq36_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	BDX_UltraSeq36_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_StdSeq50_POP7	KB_3130_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_FastSeq50_POP7	KB_3130_POP7_BDTv3direct.mob

Table 5 Run conditions for 3730/3730*xl* DNA Analyzers

Polymer	Array	Run module	Mobility file
POP-7™ polymer	36 cm	StdSeq50_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	RapidSeq36_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	TargetSeq36_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	LongSeq50_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	FastSeq50_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	BDX_StdSeq50_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	36 cm	BDX_RapidSeq36_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_LongSeq50_POP7	KB_3730_POP7_BDTv3direct.mob
POP-7™ polymer	50 cm	BDX_FastSeq50_POP7	KB_3730_POP7_BDTv3direct.mob

Appendix D – VeriDose® Panel list of genes at time of research

PHARMACOGENETICS

Variant List – VeriDose® Core Panel

Gene	Nucleotide Change	Amino Acid Change / Star Allele	dbSNP RS#
ABCB1	c.3435C>T	p.I1145=	rs1045642
APOE	c.388T>C	p.C130R	rs429358
	c.526C>T	p.R176C	rs7412
COMT	c.472G>A	p.V158M	rs4680
CYP1A2	g.-3860G>A	*1C	rs2069514
	g.-163C>A	*1F	rs762551
	g.-729C>T	*1K	rs12720461
	g.3533G>A	*7	rs56107638
	g.558C>A	*11	rs72547513
CYP2B6	c.516G>A	*6	rs3745274
	c.983T>C	*18	rs28399499
CYP2C19	c.681G>A	*2	rs4244285
	c.636G>A	*3	rs4986893
	c.1A>G	*4	rs28399504
	c.1297C>T	*5	rs56337013
	c.395G>A	*6	rs72552267
	g.19294T>A	*7	rs72558186
	c.358T>C	*8	rs41291556
	g.-806C>T	*17	rs12248560
CYP2C9	c.430C>T	*2	rs1799853
	c.1075A>C	*3	rs1057910
	c. 1076T>C	*4	rs56165452
	c.1080C>G	*5	rs28371686
	c.818delA	*6	rs9332131
	c.449G>A	*8	rs7900194
	c.1003C>T	*11	rs28371685
	c.1465C>T	*12	rs9332239
	c.269T>C	*13	rs72558187
	c.485C>A	*15	rs72558190
CYP2D6	g.2850C>T	*2, multiple	rs16947
	g.4180G>C	*2, multiple	rs1135840
	g.2549delA	*3	rs35742686
	g.1846G>A	*4	rs3892097
	g.1707delT	*6	rs5030655
	g.2935A>C	*7	rs5030867
	g.1758G>A	*8	rs5030865
	g.2613_2615delAGA (g.2615_2617delAAG)	*9	rs5030656
	g.100C>T	*10, multiple	rs1065852

For Research Use Only. Not for use in diagnostic procedures.

Agena BIOSCIENCE

Variant List – VeriDose® Panel (PAGE 2 OF 2)

Gene	Nucleotide Change	Amino Acid Change / Star Allele	dbSNP RS#
CYP2D6	c.883G>C	*11	rs201377835 rs5030863
	g.124G>A	*12	rs5030862
	g.1758G>T	*14	rs5030865
	g.137-138insT	*15	rs72549357
	g.1023C>T	*17	rs28371706
	c.dup4125-4133	*18	dup4125_4133
	c.2539_2542delAACT	*19	rs72549353
	c.1973_1974insG	*20	rs72549354
	g.3183G>A	*29	rs59421388
	g.2988G>A	*41	rs28371725
		CNV	CNV
CYP3A4	c.664T>C	*2	rs55785340
	c.566 T>C	*17	rs4987161
	g.15389C>T	*22	rs35599367
CYP3A5	g.27289C>A	*2	rs28365083
	g.6986A>G	*3	rs776746
	c.624G>A	*6	rs10264272
	g.27131_27132insT	*7	rs41303343
DRD2	c.2137G>A	p.E713K / Taq1A	rs1800497
F2	c.*97G>A	p.G20210A	rs1799963
F5	c.1601G>A	p.R506Q	rs6025
GLP1R	c.780A>C/T	p.L260F	rs1042044
	c.502G>A/C	p.G168R/G168S	rs6923761
	c.510-1135T>G		rs2300615
MTHFR	c.665C>T	p.A222V	rs1801133
	c.1286A>C	p.E429A	rs1801131
OPRM1	c.118A>G	p.A118G	rs1799971
PNPLA5	c.608-169G>A		rs5764010
SLCO1B1	c.521T>C	p.V174A	rs4149056
SULT4A1	c.743-374A>G		rs763120
VKORC1	c.-1639G>A	1639 G>A	rs9923231

Appendix E – Condensed description of recruited participant phenotypes recorded by clinicians at RCWMCH and recorded in the REDCap database

Patient Identifier	Sex	Ancestry	Age-of-Onset (Months)	Condensed Description of Phenotype
1	Female	Mixed Ancestry	72 months	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Patient described as "becoming strange", displays incoherent speech and becomes afraid, twitches left eye, hand and leg, becomes drowsy after seizures. Seizures triggered by fever and anxiety attacks. *Seizure Evolution:* No specific changes observed since onset. *Development:* Currently in mainstream school but struggling to keep up – no evidence of neuroregression. *EEG:* Abnormality described as excess slowing mixed frequencies (7-8hz). *Imaging:* CT and MRI – calcification observed. **Working Diagnosis: Focal epilepsy related to right frontoparietal calcification**
2	Male	Black African	112 months	*Current Seizure Frequency:* Once a day *Current Seizure Morphology:* Focal onset *Seizure Evolution:* Right sided focal onset, unaware, seizures sometimes tonic and generalised. Seizures were very frequent and sometimes clustered; seizures are shorter now. Seizures documented as reducing to a few a month at a few visits, but more often occur daily with no consistent pattern of increase or decrease. Seizures have clustered before with upper respiratory tract infection.

				Development: No developmental delay before seizure onset. Currently learning difficulties and cognitive delay that is global and particularly in working memory. Poor attention, slow processing speed, and difficulties with complexity and abstract thinking. Level of 4-6yr old at aged 12. Evidence of neuroregression. normal development prior to seizures with cognitive regression following onset. *EEG:* EEG background normal. *Imaging:* CT scan and MRI. No abnormalities on scans. **Working Diagnosis: Medically refractory epilepsy with regression**
3	Male	Black African	60 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Right hand gets stiff – tonic-clonic seizures. *Seizure Evolution:* Focal onset with secondary generalisation. *Development:* No developmental delay prior to seizure onset. Early development normal – global developmental delay since seizures, which appears to be getting worse. *EEG:* Symmetrical high amplitude delta transients with occipito-temporal theta activity - 3-4hz and 2-3hz transients on 2nd EEG. *Imaging:* MRI – no abnormalities on scan. **Working Diagnosis: Complex seizure disorder**
4	Female	Mixed Ancestry	3 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Goes blue and stiff, eyes roll to one side. *Seizure Evolution:* Right focal seizures with secondary generalisation - tonic seizures often nocturnal, some absence and myoclonic seizures. *Development:* No evidence of neuroregression.

				EEG: Background normal. *Imaging:* CT and MRI scans – focal cortical dysplasia observed on scan - left subependymal nodules and left pachygyria/polymicrogyria. **Working Diagnosis: Epilepsy secondary to migrational brain abnormality**
5	Male	Black African	48 months	*Current Seizure Frequency:* Once a week. *Current Seizure Morphology:* Generalised tonic-clonic seizures. Seizures trigged by fever or illness. *Seizure Evolution:* Initial generalised tonic-clonic seizures evolved into myoclonic seizures about a year later and absences noted about 6 months later. Astatic seizures/drops noted a year after that. *Development:* Moderate global developmental delay. Evidence of neuroregression - developmental delay began after onset of epilepsy and accelerated as seizure disorder evolved. *EEG:* Symmetrical 4-5hz theta activity with intermittent 1-3Hz transients. *Imaging:* CT and MRI scans – global atrophy observed on scans. **Working Diagnosis: Doose syndrome**
6	Female	Black African	96 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Generalised, febrile, and focal seizures all observed. Becomes restless, eyes and head deviate to left then left arm and leg become stiff. *Seizure Evolution:* Generalised tonic-clonic seizures, first seizure was myoclonic seizure *Development:* Developmental delay prior to seizure onset - details not provided but thought in in early years to have HIV related encephalopathy with global delay. Was slow to learn to walk.

				EEG: Bifrontal focal slowing. *Imaging:* CT and MRI scans - right temporoparietal encephalomalacia suggestive of previous vascular insult. **Working Diagnosis: Epilepsy likely secondary to HIV encephalopathy with vascular insult**
7	Female	Black African	48 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Right sided focal seizures preceded by aura and associated with loss of awareness. Fever triggers seizures. *Seizure Evolution:* Initially just febrile seizures, progressed to current seizure morphology. *Development:* Learning problems - mild developmental delay. No evidence of neuroregression. *EEG:* No abnormalities observed on EEG background. *Imaging:* CT and MRI scans – no abnormalities observed. **Working Diagnosis: Not specified in record.**
8	Male	Black African	20 months	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Myoclonic seizures with some drop attacks. *Seizure Evolution:* Recurrent febrile seizures then atypical absences - drop attacks from age 3, and now myoclonic seizures. Fever triggers seizures. *Development:* Also has cerebral palsy. Severe global developmental delay with left hemiplegia. Evidence of neuroregression - progressive falling behind as seizure disorder evolved. *EEG:* Generalised slowing on EEG background.

				Imaging: CT and MRI scans – global atrophy observed. **Working Diagnosis: Doose syndrome**
9	Male	Black African	60 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Eyes roll up, head turns to side, and jerking. Fevers and illness trigger seizures. *Seizure Evolution:* Also developed atonic drops and myoclonic jerks after generalised tonic-clonic events started - timeline not clear. *Development:* Evidence of neuroregression - had normal development early until seizures became a problem with gradual decline in intellectual function. *EEG:* Abnormal slowing with severe encephalopathy and a right posterior focus. *Imaging:* CT and MRI scans – calcification observed but MRI normal. **Working Diagnosis: Lennox-Gastaut syndrome or neurocysticercosis?**
11	Female	Black African	48 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Left focal seizures with loss of awareness. *Seizure Evolution:* Not noted in records. *Development:* Global developmental delay. Evidence of neuroregression - normal early development with some slowing of skill acquisition since epilepsy but no loss of skills noted. *EEG: S*low for age - 3-4Hz. *Imaging:* CT and MRI scans – MRI was normal, but possible signs of hypoxic-ischaemic encephalopathy on CT. **Working Diagnosis: Refractory undiagnosed focal epilepsy.**

12	Female	Black African	3 months	*Current Seizure Frequency:* Once a day. *Current Seizure Morphology:* Generalised clonic. Also occasional jerks and possibly atypical absences. *Seizure Evolution:* Initial general tonic-clonic, then myoclonic and drop attacks and then atypical absence and some focal unaware seizures. *Development:* Right hemiplegia, developmental delay before onset of seizures, was less responsive than would be expected. Evidence for neuroregression - initial mild global developmental delay evolved into severe global development delay as seizure disorder progressed. *EEG:* Symmetrical mix of frequencies with runs of sharply contoured activity. *Imaging:* CT and MRI scans – evidence of HIE, calcification, focal cortical dysplasia - left parafalcine region. **Working Diagnosis: Early Onset Epileptic Encephalopathy**
13	Male	Black African	112 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Myoclonic seizures with atypical absences but more recently focal tonic left sided seizures. *Seizure Evolution:* Initially only in sleep - myoclonic unaware. Recent focal left sided tonic events. *Development:* Had bacterial meningitis hydrocephalus with ventriculoperitoneal shunt. Post meningitis development was delayed - normal before. Has hemiplegia. Currently has severe intellectual disability with left hemiplegia and autism spectrum features. Evidence for neuroregression - only following meningitis not since onset of epilepsy although autism features only noted subsequent to seizures. *EEG:* Generalised slowing.

				Imaging: CT scan - ventriculoperitoneal shunt, right cerebral volume loss in keeping with previous infarct. **Working Diagnosis: Epilepsy secondary to cerebral insult.**
14	Male	Black African	8 months	*Current Seizure Frequency:* 5-9 a day. *Current Seizure Morphology:* Atypical absences, myoclonic jerks, left focal seizures - 2-3 times a month longer focal seizures. Fever triggers seizures. *Seizure Evolution:* Initially febrile then general tonic-clonic, evolved into left focal seizures with myoclonic jerks, atypical absences, and atonic seizures. *Development:* Evidence of neuroregression - progressively bigger gap between age related and actual function as epilepsy disorder progressed. *EEG:* Slow background. *Imaging:* CT and MRI scans – no abnormalities. **Working Diagnosis: Lennox-Gastaut Syndrome**
16	Male	Black African	36 months	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Left focal seizures, occasional myoclonic jerks. *Seizure Evolution:* Initial left sided seizures followed a few months later by some GTC seizures, evolved to include drop attacks and myoclonic jerks. *Development:* Left hemiplegia and VII palsy. Evidence of neuroregression - normal early development - speech steadily lost over a few years. *EEG:* No abnormalities described. *Imaging:* CT and MRI scans - right parietal and occipital atrophy corresponding to right sided PCA infarct. Left cerebral hemisphere mild atrophy features suggest acute chronic vascular insult.

				Working Diagnosis: Possible Lennox-Gastaut Syndrome secondary to vascular insult.
18	Male	Black African	69 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Myoclonic seizures – body "jerks". Seizures triggered by infection and fever. *Seizure Evolution:* Initially general tonic-clonic with right sided focal onset and myoclonic seizures, now mostly myoclonic seizures. *Development:* Evidence of neuroregression - early development apparently normal but function got worse with evolving seizure disorder. *EEG:* No background abnormality described in records. *Imaging:* CT and MRI imaging – possibly focal area of cystic encephalomalacia left-mid-parietal lobe on MRI; old infarct observed on CT - hypodense region in left frontoparietal area. **Working Diagnosis: Not specified in record.**
19	Female	Mixed Ancestry	8 months	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Tonic, clonic, and focal seizures, start mostly during sleep. Seizures triggered by fever. *Seizure Evolution:* Admitted with GTC but history obtained of spasms which evolved into multiple seizures - tonic, myoclonic, GTC and drop attacks – initial diagnosis of Lennox-Gastaut. *Development:* Evidence of neuroregression - as epilepsy syndrome evolved developmental delay got worse. *EEG:* No abnormalities described. *Imaging:* CT and MRI scans – global atrophy observed.

				Working Diagnosis: Early Onset Epileptic Encephalopathy
20	Female	Mixed Ancestry	15 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Right sided focal seizures - hand and face; sometimes aware and sometimes unaware, with absences and loss-of-consciousness too. Seizures triggered by fever and illness. *Seizure Evolution:* Not described in record. *Development:* Hemiplegic cerebral palsy; developmental delay before seizure onset – born extremely premature and displayed delayed development from birth. *EEG:* No abnormalities described. *Imaging:* CT and MRI scans - left ventriculomegaly ex vacuo dilatation. **Working Diagnosis: Focal seizures secondary to brain haemorrhage.**
21	Female	Black African	16 months	*Current Seizure Frequency:* A few times a week *Current Seizure Morphology:* Generalised tonic clonic a few times a month on average - "blank episodes" associated briefly with head drop more often, a few times a week. *Seizure Evolution:* 2005 – generalised tonic clonic; 2006 – tonic; 2007 – myoclonic seizures and drop attacks; 2007/8 - atypical absences. *Development:* Profound intellectual disability; evidence for neuroregression - early development not concerning but stagnated for around 3 years - never walked but was sitting and crawling at 8 months and 15 months. May have lost a few skills. Recorded as having more single words then fewer words. *EEG:* Mobile EEG showed generalised slowing and suppression but earlier EEG during sleep taken same day was normal. A follow up awake EEG was normal. *Imaging:* CT and MRI scans – showed global atrophy.

				Working Diagnosis: Lennox-Gastaut Syndrome
22	Male	European	5 months	*Current Seizure Frequency:* A few times a week *Current Seizure Morphology:* Prolonged focal unaware seizures, generalised tonic clonic seizures, myoclonic jerks. Seizures triggered by being tired, warm, or excited. *Seizure Evolution:* Initially focal right sided seizures with secondary generalisation, then evolved to multiple seizure types. *Development:* Mild to moderate learning disability; at LSEN school. Evidence of neuroregression - early development was normal – has displayed regression. *EEG:* Occasional spindle formations and V waves in sleep. *Imaging:* MRI – no abnormalities observed. **Working Diagnosis: Dravet Syndrome**
24	Female	Indian	48 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* At night patient snores and become confused, but unclear if seizures. Seizures triggered by sleep? Seizures also triggered by fever and strong emotions. *Seizure Evolution:* Febrile seizures followed by myoclonic jerks; some generalised tonic clonic seizures and then occasional drop attacks with infrequent generalised tonic clonic seizures related to fever. *Development:* Learning problems; evidence for neuroregression - no developmental concerns prior to seizures - maybe worsening developmental delay. *EEG:* No abnormalities described. *Imaging:* None. **Working Diagnosis: Doose Syndrome**

25	Female	Mixed Ancestry	114 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Right sided focal seizures starting at right leg to right hand, all stiff. Sometimes generalised tonic clonic with secondary generalisation, or not, as onset not always witnessed. *Seizure Evolution:* Similar pattern of seizures throughout. *Development:* In Grade 10, and coping. *EEG:* No abnormalities described. *Imaging:* CT and MRI scans - hypodensity suspected on CT but subsequent MRI normal. **Working Diagnosis: Focal Motor Seizures.**
26	Male	Black African	8 months	*Current Seizure Frequency:* Once a day. *Current Seizure Morphology:* Falls to the left, becomes stiff, then wakes up. *Seizure Evolution:* History of infantile spasms; then myoclonic jerks, atonic seizures, and focal tonic seizures. *Development:* Cerebral palsy with right hemiplegia; possible developmental delay prior to seizure onset - possibly spasms about 8 months, so unsure as epilepsy only diagnosed at 7 months of age. Evidence for neuroregression - seems developmental status may have worsened after seizure onset. *EEG:* No abnormalities described. *Imaging:* MRI scan - ex vacuo dilatation of left ventricle. Old lacunar infarct in left basal ganglia. **Working Diagnosis: Lennox-Gastaut Syndrome possibly secondary to infarct**

27	Male	Black African	7 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Right sided focal, unaware seizures. Seizures triggered by fever. *Seizure Evolution:* Initially generalised tonic clonic but evolved to be right sided with loss of awareness; still some generalised tonic clonic with fever. *Development:* Left Hemimegaloencephaly; severe intellectual disability; dysmorphology – epicanthic folds. *EEG:* No described abnormalities. *Imaging:* MRI scan – focal cortical dysplasia - left hemimegaloencephaly. **Working Diagnosis: Hypomelanosis of LTO with hemimegaloencephaly**
28	Male	Black African	24 months	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Drop attacks, atypical absences, some generalised tonic clonic seizures. *Seizure Evolution:* Not described in record. *Development:* Scaphocephaly or synostosis? Initial normal motor development but presented at age 2 with developmental delay and then a short time afterwards seizure noted. Evidence of neuroregression - initially seemed just developmentally delayed but seems has been a regressive element as seizure disorder has evolved. Dysmorphology – scaphocephaly noted. *EEG:* Slow. *Imaging:* MRI scan – no abnormalities noted. **Working Diagnosis: Lennox-Gastaut Syndrome**
29	Female	Black African	60 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Eyes roll back, clonic jerks, drowsy postictally.

				Seizure Evolution: Not described in records. *Development:* Development delayed prior to seizure onset – speech delayed, but no motor delay. Developmentally delayed, attends K1 LSEN. *EEG:* Awake high voltage 2-3hz delta – sleep admixture theta and delta. *Imaging:* CT and MRI scans - Ill-defined hyperintensity in deep white matter with cerebral shrinkage. **Working Diagnosis: Refractory epilepsy.**
30	Male	Black African	7 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Focal left sided seizures - turns to side then eyes up with stiffening limbs. *Seizure Evolution:* Always had a focal element - initially right then left but more recently seizures start on left. *Development:* Previously had TB meningitis; does not talk but mumbles. Cannot sit. Evidence of neuroregression. *EEG:* Generalised slowing. *Imaging:* CT scan - hydrocephalus and cerebral swelling and 2 high density foci along falx – tuberculoma? Subsequently bilateral hemisphere tuberculoma. Subsequently bilateral hemisphere and right basal ganglia infarcts. **Working Diagnosis: Epilepsy following TB meningitis.**
31	Female	Black African	72 months	*Current Seizure Frequency:* Once a day. *Current Seizure Morphology:* Left sided clonic jerks followed by post ictal sleeping. *Seizure Evolution:* No real change in semiology.

				Development: Left hemiplegia with a potential cerebral palsy cause. Dense cataract in left eye potentially since birth. All seemed fine until left sided weakness and focal seizures at 6 months of age. Has had cataract surgery and attends school for the visually impaired. *EEG:* Slow for age. *Imaging:* CT scan - right frontoparietal low density lesion with ex vacuo dilation of that ventricle and smaller right hemisphere - presumed right MCA infarct. **Working Diagnosis: Epilepsy secondary to cerebral insult**
32	Female	Black African	10 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Right sided focal and unaware seizures. Seizures triggered by fever. *Seizure Evolution:* Focal seizures and then drop attacks. *Development:* Right hemiplegia and Sturge Weber syndrome. Developmental delay observed before seizures, with mildly delayed milestones. No evidence of neuroregression. Port wine stain birthmark. *EEG:* No abnormalities described in record. *Imaging:* CT scan – evidence of HIE. Angiomatosis and left hemiatrophy, left parietal infarct on 25/04/15. **Working Diagnosis: Sturge Weber syndrome with parietal infarct**
33	Male	Mixed Ancestry	3 months	*Current Seizure Frequency:* Once a day *Current Seizure Morphology:* Myoclonic jerks *Seizure Evolution:* Infantile spasms from 3 months then episodes of staring and eye deviation from sleep at 5 years of age with abnormal EEG, then evolved into more generalised clonic

				"shaking" a year later (complex partial seizures) then myoclonic / astatic seizures described at 7 years of age with tonic posturing of arms. *Development:* Moderate developmental delay; evidence of neuroregression – displayed normal development prior to onset of spasms. *EEG:* Hypsarrhythmia. Evolved to bitemporal slowing more over left temporal area on 2nd telemetry, mostly symmetrical ay 8hz. Spindle formations and V wave recorded right centrotemporal area, frontal and left occipital in sleep more generalised spike wave or polyspike. *Imaging:* CT and MRI scans – focal cortical dysplasia – right temporoparietal dysplasia. **Working Diagnosis: Early Onset Epileptic Encephalopathy**
34	Male	Mixed Ancestry	2 months	*Current Seizure Frequency:* 5-9 a day. *Current Seizure Morphology:* About 3 myoclonic jerks that are short lived per day, 3 generalised tonic clonic seizures in last 4 months. Seizures triggered by fever. *Seizure Evolution:* Initially GTC followed by myoclonic jerks at 6 months; right focal at 2 years of age with secondary generalisation. Atonic seizures and drop attacks listed but not age of onset. *Development:* Right hemiplegia initially, now spasticity. Developmental delay observed prior to start of seizures – was not smiling by seizure onset at 10 weeks of age. Severe intellectual disability. Has microcephaly and prominent ears. *EEG:* Bilateral independent spiking more marked over left hemisphere with isolated right hemispheric slowing. *Imaging:* CT and MRI scans – no abnormalities noted. **Working Diagnosis: Early Onset Epileptic Encephalopathy**
35	Female	Mixed Ancestry	60 months	*Current Seizure Frequency:* 5-9 a day.

				Current Seizure Morphology: Myoclonic and astatic seizures and atypical absences; occasional GTC. *Seizure Evolution:* Initially myoclonic asymmetrical seizures with atypical absences - 'drop" episodes started about 1 year later. *Development:* Learning difficulties. Evidence of neuroregression - normal development until onset of seizures then increasing inattention and inability to cope with schoolwork. *EEG:* Background became slow - bilateral mixed frequencies 3hz 3-4hz transients and 4-5Hx theta activity with overriding beta activity. *Imaging:* MRI scan – no abnormalities observed. **Working Diagnosis: Doose syndrome**
36	Female	Black African	14 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Complex absence seizures, triggered by fever. *Seizure Evolution:* Initial GTCS triggered by fever then right focal tonic seizures, atonic head drops and staring spells. *Development:* Neuronal migration disorder; developmental delay observed prior to onset of seizures, global delay noted from a few months. Had motor delay prior to seizure onset. Evidence of neuroregression - as seizure disorder progressed rate of cognitive decline and developmental delay increased. Has epicanthic folds and microcephaly. *EEG:* Became slow - no post LPH rhythms when awake. *Imaging:* CT and MRI scans - pachygyria and lissencephaly – potentially calcification. **Working Diagnosis: Lennox-Gastaut syndrome secondary to neuronal migrational abnormality**
41	Female	Black African	12 months	*Current Seizure Frequency:* 1-3 per month.

				Current Seizure Morphology: Left focal seizures with loss of consciousness and residual left hemiplegia. Seizures triggered by fever. *Seizure Evolution:* Febrile GTC to left focal seizures. *Development:* Moderate global developmental delay. Evidence of neuroregression - normal early development with slowing of development as seizure disorder progressed. *EEG:* On video background somewhat attenuated with right hemisphere dominated by almost continuous PLED spike waves, max from right frontocentral regions at 2 HZ maximum in sleep, left hemisphere abnormal, poorly formed, and slow in keeping with encephalopathic state. *Imaging:* CT and MRI scans - 1st MRI normal ,2nd subtle swelling hemisphere with loss of sulcal markings. Area right high parietal region showing increase signal on FLAIR. SPECT asymmetrical perfusion temporal region - left greater than right. **Working Diagnosis: Refractory epilepsy syndrome**
42	Male	Black African	48 months	*Current Seizure Frequency:* Once a week *Current Seizure Morphology:* Left sided focal seizures. *Seizure Evolution:* In retrospect likely always focal element with secondary generalization. *Development:* Moderate intellectual disability. *EEG:* No abnormalities described in record. *Imaging:* CT and MRI scans – no abnormalities observed. **Working Diagnosis: Focal epilepsy**
44	Male	Black African	From birth	*Current Seizure Frequency:* A few times a week. *Current Seizure Morphology:* Screams, postures, tonic seizures. Seizures triggered by fever.

				Seizure Evolution: Neonatal focal seizures with generalisation but also myoclonic jerks early on. Comment of GTC after first year. *Development:* Severe global developmental delay with spasticity and dystonia. *EEG:* Background suppressed. *Imaging:* Ultrasound, CT, and MRI scans – no abnormalities observed. **Working Diagnosis: Early Onset Epileptic Encephalopathy**
45	Female	Black African	2 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Right sided tonic, unaware. Seizures triggered by fever. *Seizure Evolution:* Initially GTC recorded, then tonic spasms. *Development:* Severe global developmental delay with hemiplegia. Evidence of neuroregression - immediate neonatal period normal but floppy and delayed on presentation with epilepsy. *EEG:* Hypsarrhythmia with burst suppression. *Imaging:* MRI scan - left hemimegaloencephaly with pachygyria. **Working Diagnosis: Ohtahara syndrome**
46	Female	Mixed Ancestry	126 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Prodrome headache and nausea, legs give, can evolve to GTC seizures. Seizures triggered by headaches and heat. *Seizure Evolution:* Started with abnormal experiences - seeing and hearing things with altered awareness, then shaky episodes with altered awareness, evolved to include GTC seizures. *Development:* Learning difficulties; evidence of neuroregression - was doing fine since seizures started but some difficulties subsequently with learning.

				EEG: No abnormalities described in record. *Imaging:* MRI scan – no abnormalities described. **Working Diagnosis: Multifocal seizure disorder**
47	Male	Black African	37 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Myoclonic jerks, more frequently tonic seizures with secondary generalisation. Seizures triggered by fever. *Seizure Evolution:* Not described in record. *Development:* Had meningitis at 3 years of age. Global developmental delay. Evidence of neuroregression - normal development prior to seizure onset followed by loss of skills, but static global developmental delay now. *EEG:* No abnormalities described in record. *Imaging:* CT scan – global atrophy. White matter low density suggestive of leukoencephalopathy. **Working Diagnosis: Lennox-Gastaut syndrome**
48	Female	Mixed Ancestry	35 months	*Current Seizure Frequency:* 1-3 per month. *Current Seizure Morphology:* Eyes roll, tonic seizures, becomes sleepy - preceded by screams, myoclonic jerks. Seizures triggered by fever and infection. *Seizure Evolution:* Right focal precipitated by fever followed by a number of afebrile GTC; later blank spells and then more clusters of GTC a few months after atonic drop attacks noted, and episodes of eyes and mouth flickering and jerks. *Development:* Relatively normal? No evidence of neuroregression. *EEG:* No abnormalities described in record.

				Imaging: CT and MRI scans – no abnormalities observed. **Working Diagnosis: Lennox-Gastaut syndrome**
50	Male	Black African	29 months	*Current Seizure Frequency:* 1-3 per month *Current Seizure Morphology:* Right sided twitching of face and right side with loss of continence and becomes unaware. Seizures triggered by fever. *Seizure Evolution:* Generalised tonic clonic seizures associated with fever. *Development:* Learning difficulties. *EEG:* No abnormalities described in record. *Imaging:* MRI scan – evidence of HIE. **Working Diagnosis: Potential HIE**

Appendix F – NanoDrop readings of DNA extracted from participant samples, detailing tissue from which DNA was extracted, concentrations, and quality indicators

Family	Sample Donor	Sample Type	Concentration (ng/uL)	260/280 ratio	260/230 ratio
1	Patient	Saliva	116.46	1.81	1.17
1	Patient	Saliva	141.54	1.80	0.68
1	Parent	Saliva	71.16	1.67	0.66
1	Parent	Saliva	159.75	1.64	0.53
2	Patient	Saliva	319.32	1.97	1.65
2	Patient	Saliva	714.18	1.89	1.61
3	Patient	Saliva	100.10	1.78	1.11
3	Patient	Saliva	370.80	1.80	1.16
3	Parent	Saliva	7.77	1.35	0.30
3	Parent	Saliva	6.90	1.63	0.32
3	Parent	Saliva	4.43	1.72	0.45
3	Parent	Saliva	295.53	1.75	1.34
3	Parent	Saliva	400.84	1.82	1.24
4	Patient	Saliva	46.27	1.98	0.95
4	Parent	Saliva	215.67	1.77	1.19
4	Parent	Saliva	714.49	1.85	1.16
5	Patient	Saliva	170.97	1.87	1.06
5	Parent	Saliva	51.53	1.73	0.59
6	Patient	Saliva	1648.45	1.90	1.50
7	Patient	Saliva	329.73	1.84	1.29
7	Parent	Saliva	1367.52	1.85	1.28
7	Parent	Saliva	2069.54	1.86	1.44
8	Patient	Saliva	172.73	1.91	0.79
9	Patient	Saliva	115.59	1.87	1.06
9	Patient	Saliva	240.51	1.68	0.82
9	Parent	Saliva	963.48	1.87	1.85
11	Patient	Saliva	1194.25	1.90	1.48
11	Parent	Saliva	1443.80	1.89	1.57
12	Patient	Saliva	255.34	1.92	1.13

12	Parent	Saliva	859.24	1.84	1.42
13	Patient	Saliva	439.61	1.69	0.79
13	Parent	Saliva	870.48	1.84	1.83
13	Parent	Saliva	828.38	1.82	1.60
14	Patient	Saliva	284.07	1.89	1.34
16	Patient	Saliva	585.96	1.93	1.46
16	Parent	Saliva	1153.74	1.76	1.19
18	Patient	Saliva	297.29	1.83	1.34
19	Patient	Saliva	272.54	1.79	0.82
19	Parent	Saliva	242.37	1.74	1.04
20	Patient	Saliva	157.07	1.52	0.35
21	Patient	Saliva	260.59	1.86	0.89
21	Parent	Saliva	117.50	1.47	0.47
21	Parent	Saliva	498.48	1.81	1.46
22	Patient	Blood	719.71	1.82	1.77
24	Patient	Saliva	175.19	1.93	0.87
24	Parent	Saliva	743.98	1.89	1.66
24	Parent	Saliva	1687.25	1.86	1.55
25	Patient	Saliva	1065.33	1.85	1.57
25	Parent	Saliva	1479.96	1.79	1.24
26	Patient	Saliva	481.00	1.70	0.98
26	Parent	Saliva	836.83	1.85	1.46
27	Patient	Saliva	419.40	1.94	1.38
27	Parent	Saliva	628.12	1.79	1.24
27	Patient	Saliva	3.19	2.38	0.39
27	Patient	Saliva	23.66	2.09	2.50
28	Patient	Saliva	2644.62	1.98	1.57
28	Parent	Saliva	2209.44	1.97	1.66
29	Patient	Saliva	88.48	1.26	0.32
29	Parent	Saliva	1821.46	1.89	1.37
30	Patient	Blood	22.4	1.87	-2.78
30	Parent	Saliva	1398.84	1.93	1.69
30	Parent	Saliva	1995.22	1.85	1.44
31	Patient	Saliva	279.38	1.65	0.65
31	Parent	Saliva	1880.70	1.81	1.27

32	Patient	Saliva	484.21	1.87	1.37
32	Parent	Saliva	1440.77	1.85	1.92
33	Patient	Saliva	90.70	1.74	0.87
33	Parent	Saliva	157.74	1.53	0.46
34	Patient	Saliva	341.01	1.83	0.75
34	Parent	Saliva	1996.34	1.88	1.62
35	Patient	Saliva	277.18	1.79	0.98
35	Parent	Saliva	1091.99	2.00	1.90
35	Parent	Saliva	117.70	1.64	0.61
36	Patient	Saliva	221.37	1.95	0.91
36	Parent	Saliva	145.57	1.53	0.42
41	Patient	Saliva	128.38	1.60	0.60
41	Parent	Saliva	125.05	1.77	0.71
42	Patient	Saliva	201.53	1.76	0.85
42	Parent	Saliva	658.08	1.87	1.64
42	Parent	Saliva	1988.75	1.87	1.87
44	Patient	Blood	488.03	1.78	1.62
44	Parent	Saliva	572.44	1.71	1.27
44	Parent	Saliva	688.60	1.85	1.55
45	Patient	Saliva	326.35	1.88	1.00
45	Parent	Saliva	565.71	1.87	1.63
46	Patient	Saliva	291.74	1.63	0.62
46	Parent	Saliva	1948.94	1.83	1.45
47	Patient	Saliva	321.74	1.85	1.25
47	Parent	Saliva	425.31	1.81	1.58
48	Patient	Saliva	1156.83	1.87	1.27
48	Parent	Saliva	396.70	1.83	1.46
50	Patient	Saliva	491.72	1.86	1.06
50	Parent	Saliva	711.26	1.90	1.65

Note An abnormal 280/260 ratio may indicate the presence of proteins or phenol; an abnormal 260/230 ratio may indicate chemical contamination of the sample by organic compounds. Multiple DNA samples were sometime available from one individual, hence the disparity between the total number of participants and the total number of samples.

Appendix G – Full Run Reports exported from Torrent Suite™ Software

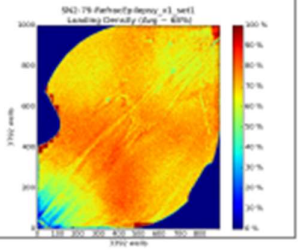
Run Report for Auto_user_SN2-79-RefracEpilepsy_v1_set1_274

Run Summary

1.21 G	77
Total Bases	Key Signal

68%
ISP Loading

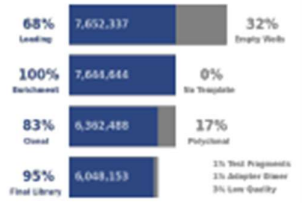
ISP Density



6,048,153
Total Reads

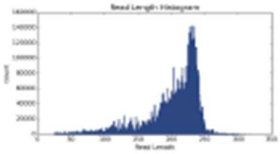
80%
Usable Reads

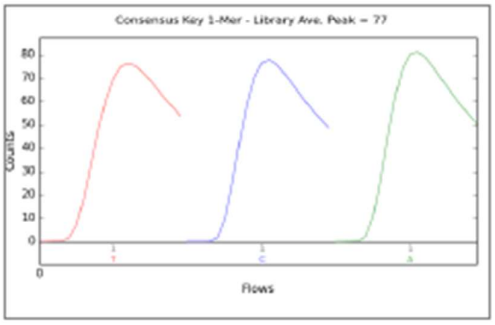
ISP Summary



200 bp	214 bp	231 bp
Mean	Median	Mode

Read Length





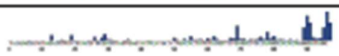
Addressable Wells	11,287,275	
With ISPs	7,652,337	67.8%
Live	7,644,644	99.9%
Test Fragment	73,236	01.0%
Library	7,571,408	99.0%
Library ISPs	**7,571,408**	
Filtered: Polyclonal	1,282,156	16.9%
Filtered: Low Quality	197,867	02.6%
Filtered: Adapter Dimer	43,232	00.6%
Final Library ISPs	**6,048,153**	**79.9%**

Barcode Name	Sample	Bases	≥ Q20	Reads	Mean Read Length	Read Length Histogram
No barcode	none	38,078,925	36,218,486	183,523	207 bp	
IonCode_0109	EE154.1NAD	155,819,488	149,478,229	782,745	199 bp	
IonCode_0110	EIEE24.1AZA	165,527,249	158,721,914	832,837	198 bp	
IonCode_0111	RE3.1ANA	173,867,632	166,742,594	880,929	197 bp	
IonCode_0112	RE7.1UNA	166,722,134	159,868,602	839,902	198 bp	
IonCode_0113	EE29.1NTA	164,376,721	157,827,034	814,827	201 bp	
IonCode_0114	RE2.1CAS	137,817,117	132,277,977	685,878	200 bp	
IonCode_0115	RE5.1GUG	85,792,653	82,346,102	425,839	201 bp	

1

iontorrent
by Thermo Fisher Scientific

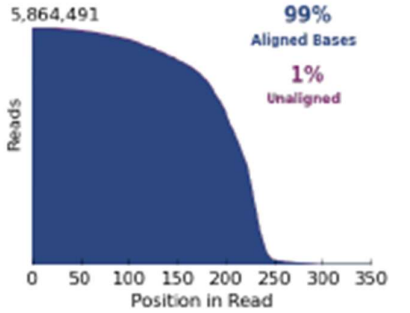
IonCode_0116	RE6.1KHA	121,510,567	116,693,271	601,534	202 bp	

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_C	**27,434**	**85**	
TF_1	**44,288**	**96**	

Alignment Summary *(aligned to Homo sapiens)*

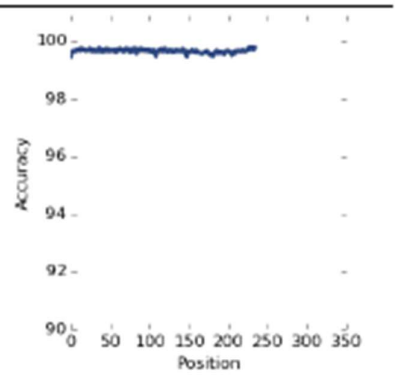
1.16 G
Total Alignment Bases

0.4X
Average Coverage Depth of Reference



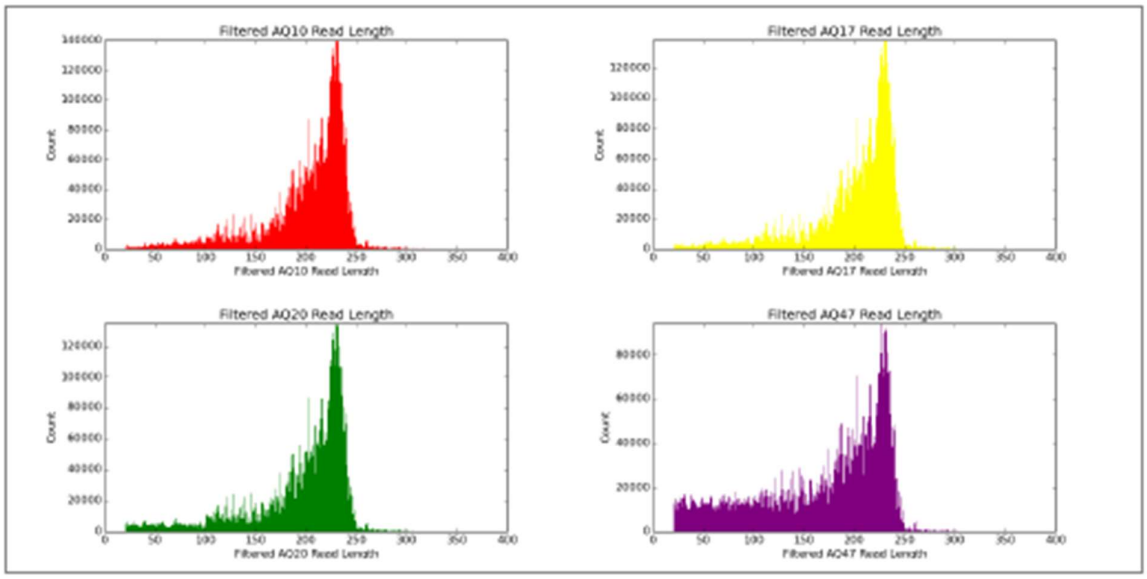
	Count	%
Total Reads	5,864,491	–
Aligned Reads	5,845,469	99.7%
Unaligned Reads	19,022	0.3%

99.7%
Mean Raw Accuracy 1x



Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.15 G	1.12 G	940 M
Mean Length [bp]	198	194	168
Longest Alignment [bp]	340	339	337
Mean Coverage Depth	0.4	0.4	0.3


Filtered AQ10 Read Length
Count
Filtered AQ10 Read Length
Filtered AQ17 Read Length
Count
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Count
Filtered AQ20 Read Length
Filtered AQ47 Read Length
Count
Filtered AQ47 Read Length

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD190273_182_RefracEpilepsySpikeIn

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonCode_0109	EE154.1NAD	780,536	96.08%	369.5	95.14%
IonCode_0110	EIEE24.1AZA	826,741	96.47%	392.9	94.12%
IonCode_0111	RE3.1ANA	877,027	94.61%	407.7	94.10%
IonCode_0112	RE7.1UNA	835,841	95.20%	391.8	94.35%
IonCode_0113	EE29.1NTA	813,292	97.02%	392.1	94.18%
IonCode_0114	RE2.1CAS	684,842	97.67%	330.9	94.60%
IonCode_0115	RE5.1GUG	424,911	97.90%	206.2	93.42%
IonCode_0116	RE6.1KHA	600,279	97.26%	290.5	94.16%

1 5 items per page 1 - 8 of 8 items

Analysis Details

Run Name	R_2015_02_20_01_40_31_user_SN2-79-RefracEpilepsy_v1_set1
Run Date	Feb. 20, 2015, 1:40 a.m.
Run Flows	500
Projects	Panels_DNA_Ampliseq
Samples	RE2.1CAS, EIEE24.1AZA, EE29.1NTA, RE3.1ANA, RE5.1GUG, RE6.1KHA, RE7.1UNA, EE154.1NAD
Reference	
Instrument	sn274670980
Operation Mode	Customer mode
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Barcode	AB0351306
Chip Check	Passed
Chip Type	318C
Chip Data	single
Chip Lot Number	Missing
Chip Wafer	Missing
Barcode Set	IonCode Barcodes 1-32
Analysis Name	Auto_user_SN2-79-RefracEpilepsy_v1_set1_274
Analysis Date	Jan. 30, 2020, 9:58 p.m.
Analysis Flows	500
runID	SX0DS
BeadFind Args	justBeadFind –args-json /opt/ion/config/args_318_beadfind.json
Analysis Args	Analysis –args-json /opt/ion/config/args_318_analysis.json –gopt /opt/ion/config/gopt_318v2_Hi-Q.param.json
Pre-BaseCaller	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20
Calibration Args	Calibration
BaseCaller Args	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20 –phred-table-file /opt/ion/config/phredTable.318.B5.h5
Alignment Args	tmap mapall ... stage1 map4
IonStats Args	ionstats alignment
Analysis Parameters	default

Chef Summary

Chef Library Prep Information for Sample Set RefracEpilepsy_set1 :

Library Prep Type	AmpliSeq on Chef
Library Prep Plate Type	BC 9-16 (Yellow)
PCR Plate Serial Number	A020018330
Combined Library Tube Label	C0574022
Last Updated	Jan. 28, 2020, 5:26 p.m.
Instrument Name	CHEF00830
Operation Mode	Customer Mode
Tip Rack Barcode	6960601C8
Kit Type	Ion AmpliSeq Kit for Chef DL8
Reagent Lot Number	2112588
Reagent Part Number	A29025C
Reagent Expiration	201130
Solution Lot Number	2115205
Solution Part Number	A29026C
Solution Expiration	201231
Script Version	905
Package Version	IC.5.12.1.RC.1
Start Time	Jan. 28, 2020, 10:31 a.m.
Completion Time	Jan. 28, 2020, 5:25 p.m.

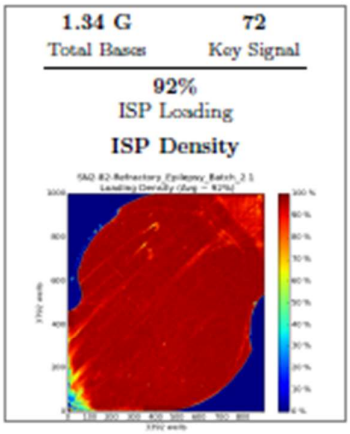
Chef Template Prep Information:

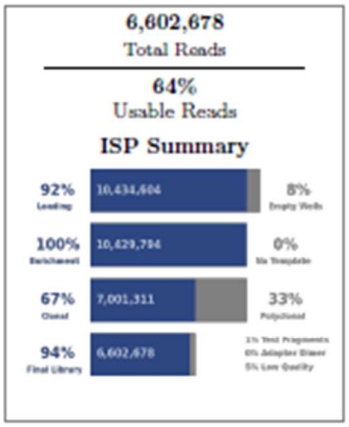
Chef Last Updated	Jan. 30, 2020, 8:35 a.m.
Chef Instrument Name	CHEF00830
Chef Operation Mode	Customer Mode
Sample Position	2
Tip Rack Barcode	48709025D
Chip Type 1	318v2
Chip Type 2	318v2
Chip Expiration 1	None
Chip Expiration 2	None
Templating Kit Type	Ion PGM Hi-Q View Chef Kit
Chef Flexible Workflow	
Reagent Expiration	191231
Reagent Lot Number	2047473
Reagent Part Number	A29901C
Reagent Cartridge Serial Number	None
Solution Lot Number	2112881
Solution Part Number	A25956C
Templating Protocol Planned	Chef Protocol - 200 bp
Solution Cartridge Serial Number	None
Solution Expiration	200531
Templating Protocol Executed	Chef Protocol - 200 bp
Chef Script Version	905
Chef Package Version	IC.5.12.1.RC.1
Start Time	Jan. 29, 2020, 4:20 p.m.
Completion Time	Jan. 30, 2020, 8:35 a.m.

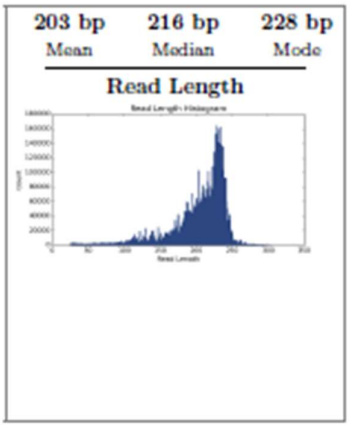
Software Version

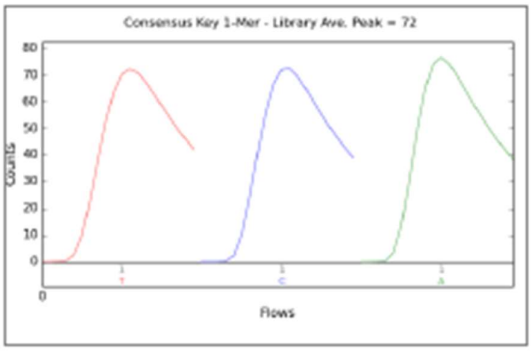
Torrent_Suite	5.12.1
host	5QYTV12
ion-analysis	5.12.27-1
ion-dbreports	5.12.60-1
ion-gpu	5.12.1-1
ion-pipeline	5.12.17-1
ion-torrentpy	5.12.21-1
ion-torrentr	5.12.23-1
Script	22.0.0
LiveView	755
DataCollect	489
OS	21
Graphics	36
Ion_Chef	IC.5.12.1.RC.1

Run Summary





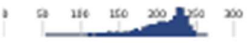


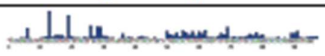


Addressable Wells	11,287,275	
With ISPs	10,434,604	92.4%
Live	10,429,794	100.0%
Test Fragment	41,724	00.4%
Library	10,388,070	99.6%
Library ISPs	**10,388,070**	
Filtered: Polyclonal	3,428,483	33.0%
Filtered: Low Quality	349,485	03.4%
Filtered: Adapter Dimer	7,424	00.1%
Final Library ISPs	**6,602,678**	**63.6%**

Barcode Name	Sample	Bases	$\geq Q20$	Reads	Mean Read Length	Read Length Histogram
No barcode	none	45,373,319	42,688,172	220,541	205 bp	
IonCode_0101	DS 4.1PHI	153,290,552	145,376,677	756,687	202 bp	
IonCode_0102	EE 14.1ASE	162,691,100	154,710,087	800,718	203 bp	
IonCode_0103	RE 8.1SIP	177,047,826	168,506,828	869,204	203 bp	
IonCode_0104	RE 19.1ROM	171,314,167	162,572,947	843,526	203 bp	
IonCode_0105	RE 23.1SIP	160,168,456	151,848,555	791,286	202 bp	
IonCode_0106	RE 25.1LIN	155,315,864	147,183,580	763,635	203 bp	
IonCode_0107	RE 12.1SIY	155,255,691	147,791,021	764,163	203 bp	

1

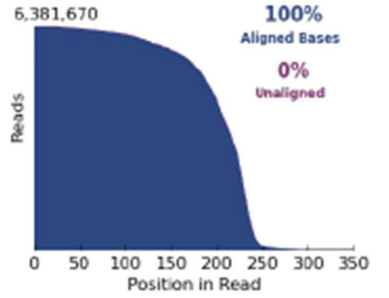
IonCode_0108	RE 14.1SIN	160,826,943	152,228,285	792,451	202 bp	

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_C	**8,491**	**51**	
TF_1	**30,849**	**95**	

Alignment Summary *(aligned to Homo sapiens)*

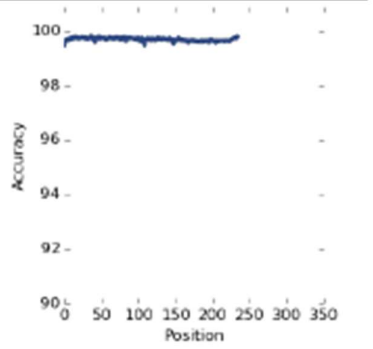
1.29 G
Total Alignment Bases

0.4X
Average Coverage Depth of Reference



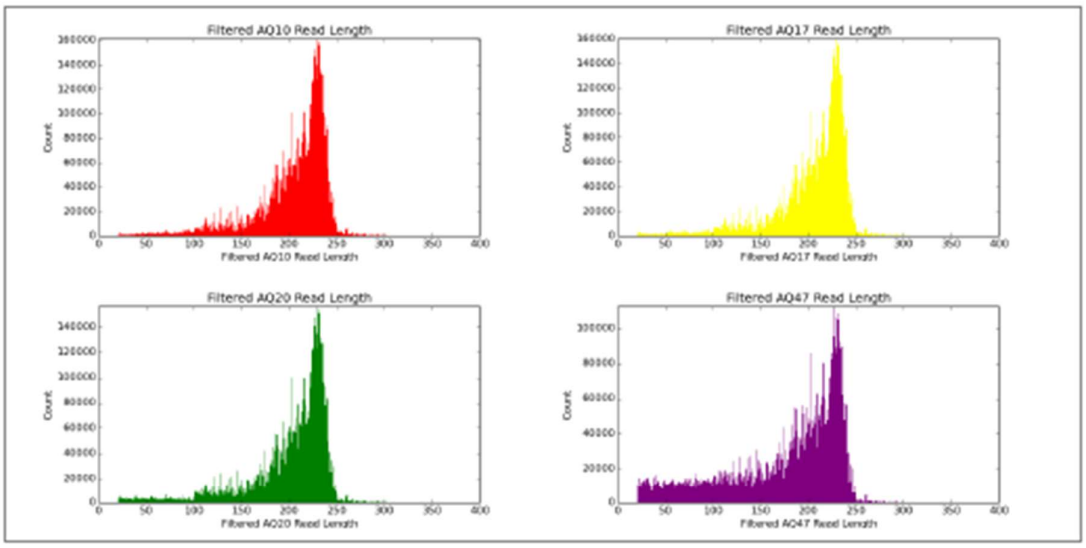
	Count	%
Total Reads	6,381,670	–
Aligned Reads	6,365,943	99.8%
Unaligned Reads	15,727	0.2%

99.7%
Mean Raw Accuracy 1x



Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.28 G	1.25 G	1.07 G
Mean Length [bp]	202	199	174
Longest Alignment [bp]	341	341	333
Mean Coverage Depth	0.4	0.4	0.3


Filtered AQ10 Read Length
Filtered AQ10 Read Length
Count
Filtered AQ17 Read Length
Filtered AQ17 Read Length
Count
Filtered AQ20 Read Length
Filtered AQ20 Read Length
Count
Filtered AQ47 Read Length
Filtered AQ47 Read Length
Count

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD190273_182_RefracEpilepsySpikeIn

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonCode_0101	DS 4.1PHI	755,070	97.76%	368.3	94.36%
IonCode_0102	EE 14.1ASE	796,752	96.09%	391.8	94.02%
IonCode_0103	RE 8.1SIP	866,955	97.85%	425.5	93.57%
IonCode_0104	RE 19.1ROM	840,963	97.00%	408.4	94.05%
IonCode_0105	RE 23.1SIP	769,321	97.63%	354.2	94.19%
IonCode_0106	RE 25.1LIN	762,264	98.30%	374.8	93.18%
IonCode_0107	RE 12.1SIY	762,084	97.63%	373.1	93.40%
IonCode_0108	RE 14.1SIN	790,534	97.38%	384.8	94.20%

1 | 5 items per page | 1 - 8 of 8 items

Analysis Details

Run Name	R_2015_02_19_18_45_26_user_SN2-82-Refractory_Epilepsy_Batch_2.1
Run Date	Feb. 19, 2015, 6:45 p.m.
Run Flows	500
Projects	
Samples	RE_25.1LIN, RE_23.1SIP, DS_4.1PHI, RE_8.1SIP, RE_12.1SIY, RE_19.1ROM, EE_14.1ASE, RE_14.1SIN
Reference	
Instrument	sn274670980
Operation Mode	Customer mode
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Barcode	AB0351303
Chip Check	Passed
Chip Type	318C
Chip Data	single
Chip Lot Number	Missing
Chip Wafer	Missing
Barcode Set	IonCode Barcodes 1-32
Analysis Name	Auto_user_SN2-82-Refractory_Epilepsy_Batch_2.1_283
Analysis Date	March 13, 2020, 5:37 p.m.
Analysis Flows	500
runID	CLMNF
BeadFind Args	justBeadFind –args-json /opt/ion/config/args_318_beadfind.json
Analysis Args	Analysis –args-json /opt/ion/config/args_318_analysis.json –gopt /opt/ion/config/gopt_318v2_Hi-Q.param.json
Pre-BaseCaller	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20
Calibration Args	Calibration
BaseCaller Args	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20 –phred-table-file /opt/ion/config/phredTable.318.B5.h5
Alignment Args	tmap mapall ... stage1 map4
IonStats Args	ionstats alignment
Analysis Parameters	default

Chef Summary

Chef Template Prep Information:

Chef Last Updated	March 13, 2020, 9:37 a.m.
Chef Instrument Name	CHEF00830
Chef Operation Mode	Customer Mode
Sample Position	1
Tip Rack Barcode	48716003E
Chip Type 1	318v2
Chip Type 2	318v2
Chip Expiration 1	None
Chip Expiration 2	None
Templating Kit Type	Ion PGM Hi-Q View Chef Kit
Chef Flexible Workflow	
Reagent Expiration	200531
Reagent Lot Number	2112884
Reagent Part Number	A29901C
Reagent Cartridge Serial Number	None
Solution Lot Number	2112881
Solution Part Number	A25956C
Templating Protocol Planned	Chef Protocol - 200 bp
Solution Cartridge Serial Number	None
Solution Expiration	200531
Templating Protocol Executed	Chef Protocol - 200 bp
Chef Script Version	905
Chef Package Version	IC.5.12.1.RC.1
Start Time	March 12, 2020, 3:47 p.m.
Completion Time	March 13, 2020, 9:37 a.m.

Software Version

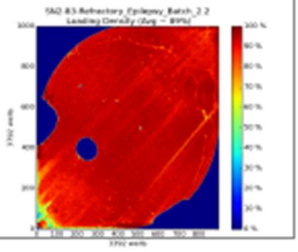
Torrent_Suite	5.12.1
host	5QYTV12
ion-analysis	5.12.27-1
ion-dbreports	5.12.60-1
ion-gpu	5.12.1-1
ion-pipeline	5.12.17-1
ion-torrentpy	5.12.21-1
ion-torrentr	5.12.23-1
Script	22.0.0
LiveView	755
DataCollect	489
OS	21
Graphics	36
Ion_Chef	IC.5.12.1.RC.1

Run Summary

1.4 G	74
Total Bases	Key Signal

89%
ISP Loading

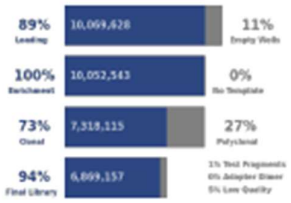
ISP Density



6,869,157
Total Reads

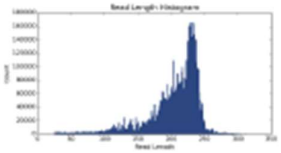
69%
Usable Reads

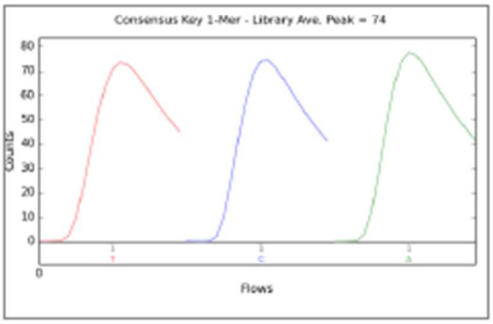
ISP Summary



204 bp	216 bp	231 bp
Mean	Median	Mode

Read Length





Addressable Wells	11,287,275	
With ISPs	10,069,628	89.2%
Live	10,052,543	99.8%
Test Fragment	57,941	00.6%
Library	9,994,602	99.4%
Library ISPs	**9,994,602**	
Filtered: Polyclonal	2,734,428	27.4%
Filtered: Low Quality	377,986	03.8%
Filtered: Adapter Dimer	13,031	00.1%
Final Library ISPs	**6,869,157**	**68.7%**

Barcode Name	Sample	Bases	≥ Q20	Reads	Mean Read Length	Read Length Histogram
No barcode	none	46,080,215	43,820,749	219,422	210 bp	
IonCode_0109	16.1ASI	174,789,558	167,239,714	864,404	202 bp	
IonCode_0110	17.1STA	192,436,879	183,852,055	943,677	203 bp	
IonCode_0111	11.1JAI	178,350,799	170,693,894	870,724	204 bp	
IonCode_0112	20.1HAN	131,273,319	125,827,637	654,917	200 bp	
IonCode_0113	21.1LIA	200,655,770	192,158,529	981,711	204 bp	
IonCode_0114	24.1OKU	154,432,229	147,919,165	757,919	203 bp	
IonCode_0115	15.1DEO	145,131,796	138,943,360	712,522	203 bp	

ion torrent
by Thermo Fisher Scientific

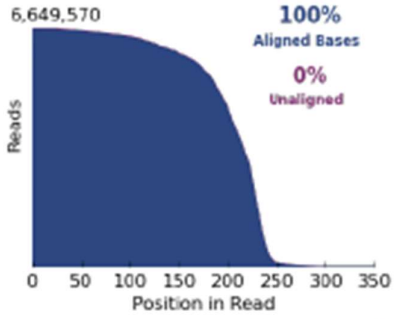
IonCode_0116	9.1SIH	175,054,290	167,601,866	863,696	202 bp	

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_C	**14,363**	**65**	
TF_1	**40,842**	**96**	

Alignment Summary *(aligned to Homo sapiens)*

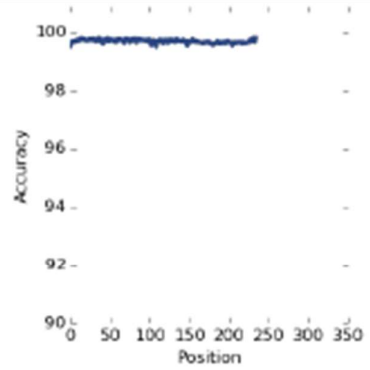
1.35 G
Total Alignment Bases

0.4X
Average Coverage Depth of Reference



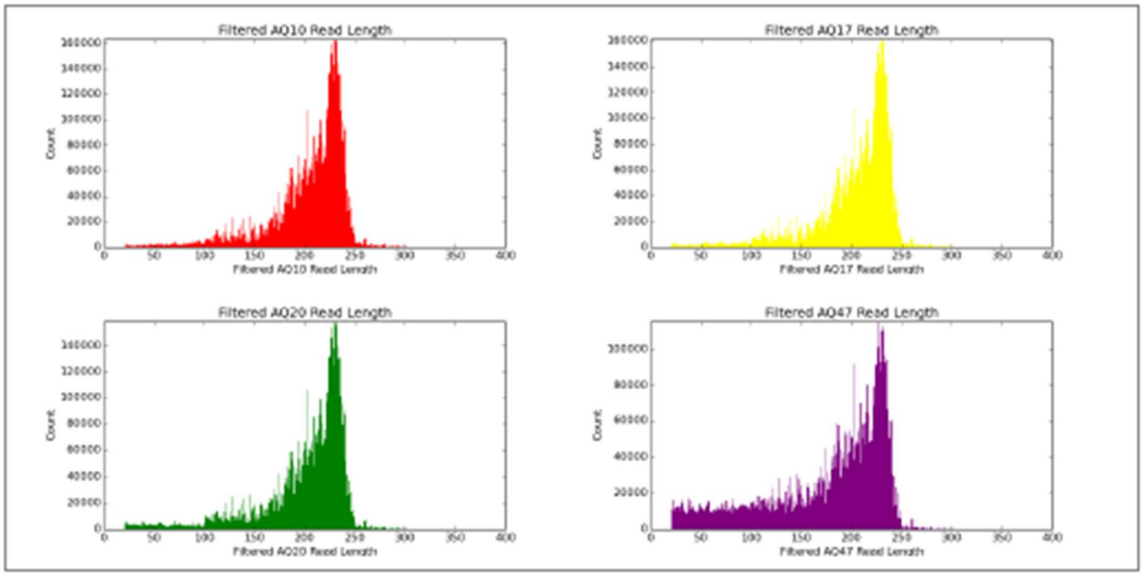
	Count	%
Total Reads	6,649,570	–
Aligned Reads	6,631,446	99.7%
Unaligned Reads	18,124	0.3%

99.7%
Mean Raw Accuracy 1x



Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.33 G	1.3 G	1.12 G
Mean Length [bp]	202	199	175
Longest Alignment [bp]	346	342	332
Mean Coverage Depth	0.4	0.4	0.4


Filtered AQ10 Read Length
Count
Filtered AQ10 Read Length
Filtered AQ17 Read Length
Count
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Count
Filtered AQ20 Read Length
Filtered AQ47 Read Length
Count
Filtered AQ47 Read Length

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD190273_182_RefracEpilepsySpikeIn

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonCode_0109	16.1ASI	862,119	97.38%	418.7	94.90%
IonCode_0110	17.1STA	941,140	97.92%	462.6	94.07%
IonCode_0111	11.1JAI	868,177	97.86%	428.6	93.96%
IonCode_0112	20.1HAN	653,769	97.72%	315.8	92.18%
IonCode_0113	21.1LIA	979,035	97.48%	480.5	93.66%
IonCode_0114	24.1OKU	755,726	97.57%	370.3	94.91%
IonCode_0115	15.1DEO	711,306	97.80%	348.6	93.69%
IonCode_0116	9.1SIH	860,174	97.93%	420.5	94.46%

1 5 items per page 1 - 8 of 8 items

Analysis Details

Run Name	R_2015_02_19_23_21_24_user_SN2-83-Refractory_Epilepsy_Batch_2.2
Run Date	Feb. 19, 2015, 11:21 p.m.
Run Flows	500
Projects	
Samples	17.1STA, 11.1JAI, 24.1OKU, 20.1HAN, 9.1SIH, 16.1ASI, 15.1DEO, 21.1LIA
Reference	
Instrument	sn274670980
Operation Mode	Customer mode
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Barcode	AB0351324
Chip Check	Passed
Chip Type	318C
Chip Data	single
Chip Lot Number	Missing
Chip Wafer	Missing
Barcode Set	IonCode Barcodes 1-32
Analysis Name	Auto_user_SN2-83-Refractory_Epilepsy_Batch_2.2_282
Analysis Date	March 13, 2020, 9:21 p.m.
Analysis Flows	500
runID	WJ332
BeadFind Args	justBeadFind –args-json /opt/ion/config/args_318_beadfind.json
Analysis Args	Analysis –args-json /opt/ion/config/args_318_analysis.json –gopt /opt/ion/config/gopt_318v2_Hi-Q.param.json
Pre-BaseCaller	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20
Calibration Args	Calibration
BaseCaller Args	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20 –phred-table-file /opt/ion/config/phredTable.318.B5.h5
Alignment Args	tmap mapall ... stage1 map4
IonStats Args	ionstats alignment
Analysis Parameters	default

Chef Summary

Chef Library Prep Information for Sample Set Epilepsy Batch 2.2:

Library Prep Type	AmpliSeq on Chef
Library Prep Plate Type	BC 9-16 (Yellow)
PCR Plate Serial Number	A020021085
Combined Library Tube Label	C0627109
Last Updated	March 12, 2020, 3:07 p.m.
Instrument Name	CHEF00830
Operation Mode	Customer Mode
Tip Rack Barcode	69827024D
Kit Type	Ion AmpliSeq Kit for Chef DL8
Reagent Lot Number	2132447
Reagent Part Number	A29025C
Reagent Expiration	201231
Solution Lot Number	2115205
Solution Part Number	A29026C
Solution Expiration	201231
Script Version	905
Package Version	IC.5.12.1.RC.1
Start Time	March 12, 2020, 8:11 a.m.
Completion Time	March 12, 2020, 3:07 p.m.

Chef Template Prep Information:

Chef Last Updated	March 13, 2020, 9:37 a.m.
Chef Instrument Name	CHEF00830
Chef Operation Mode	Customer Mode
Sample Position	2
Tip Rack Barcode	48716003E
Chip Type 1	318v2
Chip Type 2	318v2
Chip Expiration 1	None
Chip Expiration 2	None
Templating Kit Type	Ion PGM Hi-Q View Chef Kit
Chef Flexible Workflow	
Reagent Expiration	200531
Reagent Lot Number	2112884
Reagent Part Number	A29901C
Reagent Cartridge Serial Number	None
Solution Lot Number	2112881
Solution Part Number	A25956C
Templating Protocol Planned	Chef Protocol - 200 bp
Solution Cartridge Serial Number	None
Solution Expiration	200531
Templating Protocol Executed	Chef Protocol - 200 bp
Chef Script Version	905
Chef Package Version	IC.5.12.1.RC.1
Start Time	March 12, 2020, 3:47 p.m.
Completion Time	March 13, 2020, 9:37 a.m.

Software Version

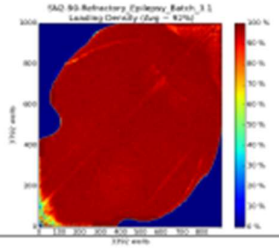
Torrent_Suite	5.12.1
host	5QYTV12
ion-analysis	5.12.27-1
ion-dbreports	5.12.60-1
ion-gpu	5.12.1-1
ion-pipeline	5.12.17-1
ion-torrentpy	5.12.21-1
ion-torrentr	5.12.23-1
Script	22.0.0
LiveView	755
DataCollect	489
OS	21
Graphics	36
Ion_Chef	IC.5.12.1.RC.1

Run Summary

1.12 G Total Bases | **71** Key Signal

92% ISP Loading

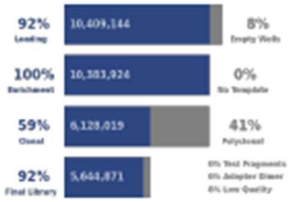
ISP Density



5,644,871 Total Reads

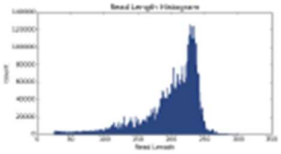
54% Usable Reads

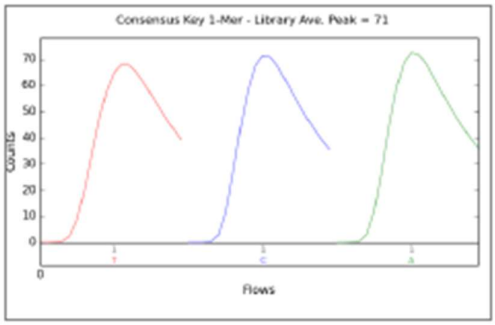
ISP Summary



198 bp Mean | **212 bp** Median | **228 bp** Mode

Read Length





Addressable Wells	**11,287,275**	
With ISPs	10,409,144	92.2%
Live	10,383,924	99.8%
Test Fragment	14,132	00.1%
Library	10,369,792	99.9%
Library ISPs	**10,369,792**	
Filtered: Polyclonal	4,255,905	41.0%
Filtered: Low Quality	459,612	04.4%
Filtered: Adapter Dimer	9,404	00.1%
Final Library ISPs	**5,644,871**	**54.4%**

Barcode Name	Sample	Bases	≥ Q20	Reads	Mean Read Length	Read Length Histogram
No barcode	none	43,719,372	39,391,713	221,332	197 bp	
IonCode_0117	RE13.1ZIM	137,617,075	125,934,362	691,785	198 bp	
IonCode_0118	RE18.1BUK	134,897,584	123,000,904	683,102	197 bp	
IonCode_0119	RE21.1LIS	121,541,399	111,243,654	613,297	198 bp	
IonCode_0120	RE22.1IMA	145,478,306	132,462,105	735,873	197 bp	
IonCode_0121	RE26.1IMA	135,279,191	123,402,754	686,033	197 bp	
IonCode_0122	RE27.1NJA	124,037,015	113,717,370	623,270	199 bp	
IonCode_0123	RE28.1JAD	136,133,484	124,812,366	688,226	197 bp	

1

ion torrent by Thermo Fisher Scientific

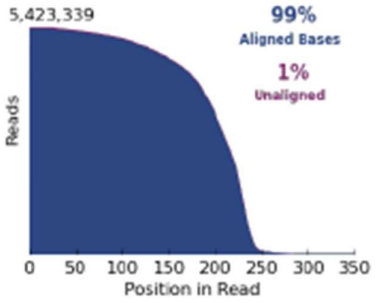
IonCode_0124	RE29.1SIS	139,505,746	127,763,653	701,753	198 bp	

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_1	**11,723**	**88**	

Alignment Summary *(aligned to Homo sapiens)*

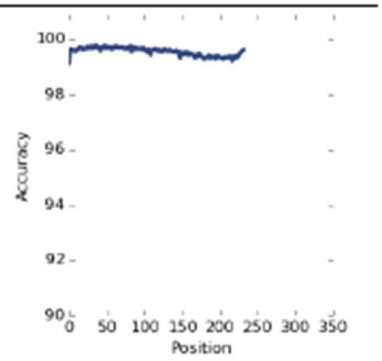
1.07 G
Total Alignment Bases

0.3X
Average Coverage Depth of Reference



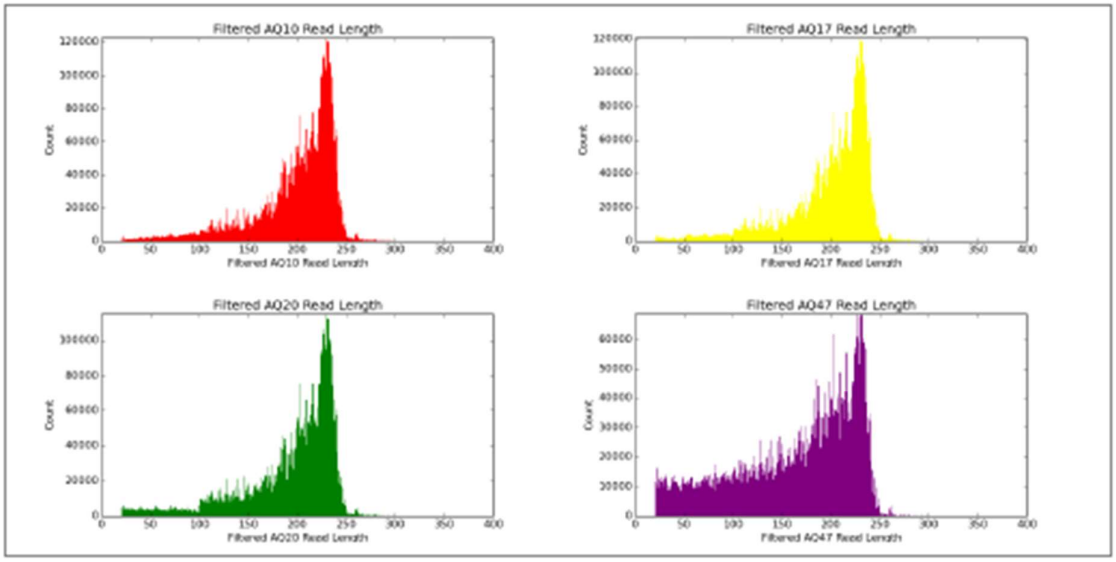
	Count	%
Total Reads	5,423,339	–
Aligned Reads	5,387,508	99.3%
Unaligned Reads	35,831	0.7%

99.6%
Mean Raw Accuracy 1x



Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.05 G	1.01 G	836 M
Mean Length [bp]	197	192	164
Longest Alignment [bp]	342	340	340
Mean Coverage Depth	0.3	0.3	0.3


Filtered AQ10 Read Length
Count
Filtered AQ10 Read Length
Filtered AQ17 Read Length
Count
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Count
Filtered AQ20 Read Length
Filtered AQ47 Read Length
Count
Filtered AQ47 Read Length

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD190273_182_RefracEpilepsySpikeIn

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonCode_0117	RE13.1ZIM	683,755	97.56%	327.6	92.44%
IonCode_0118	RE18.1BUK	650,194	97.67%	322.9	91.66%
IonCode_0119	RE21.1LIS	610,040	97.74%	291.1	91.22%
IonCode_0120	RE22.1IMA	732,133	97.67%	348.1	94.18%
IonCode_0121	RE26.1IMA	679,932	97.88%	323.1	93.66%
IonCode_0122	RE27.1NJA	619,540	97.92%	297.4	90.18%
IonCode_0123	RE28.1JAD	682,474	97.77%	325.3	90.95%
IonCode_0124	RE29.1SIS	699,440	98.15%	335.6	92.38%

1 5 items per page 1 - 8 of 8 items

ion torrent by Thermo Fisher Scientific

Analysis Details

Run Name	R_2015_02_19_19_04_48_user_SN2-90-Refractory_Epilepsy_Batch_3.1
Run Date	Feb. 19, 2015, 7:04 p.m.
Run Flows	500
Projects	
Samples	RE27.1NJA, RE18.1BUK, RE13.1ZIM, RE29.1SIS, RE21.1LIS, RE28.1JAD, RE26.1IMA, RE22.1IMA
Reference	
Instrument	sn274670980
Operation Mode	Customer mode
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Barcode	AB0357453
Chip Check	Passed
Chip Type	318C
Chip Data	single
Chip Lot Number	Missing
Chip Wafer	Missing
Barcode Set	IonCode Barcodes 1-32
Analysis Name	Auto_user_SN2-90-Refractory_Epilepsy_Batch_3.1_290
Analysis Date	July 23, 2020, 5:50 p.m.
Analysis Flows	500
runID	BWVEX
BeadFind Args	justBeadFind –args-json /opt/ion/config/args_318_beadfind.json
Analysis Args	Analysis –args-json /opt/ion/config/args_318_analysis.json –gopt /opt/ion/config/gopt_318v2_Hi-Q.param.json
Pre-BaseCaller	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20
Calibration Args	Calibration
BaseCaller Args	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20 –phred-table-file /opt/ion/config/phredTable.318.B5.h5
Alignment Args	tmap mapall ... stage1 map4
IonStats Args	ionstats alignment
Analysis Parameters	default

Chef Summary

Chef Library Prep Information for Sample Set Epilepsy Batch 3.1:

Library Prep Type	AmpliSeq on Chef
Library Prep Plate Type	BC 17-24 (Green)
PCR Plate Serial Number	A030019948
Combined Library Tube Label	C0681325
Last Updated	July 21, 2020, 5:55 p.m.
Instrument Name	CHEF00830
Operation Mode	Customer Mode
Tip Rack Barcode	69925018C
Kit Type	Ion AmpliSeq Kit for Chef DL8
Reagent Lot Number	2189459
Reagent Part Number	A29025C
Reagent Expiration	210331
Solution Lot Number	2144926
Solution Part Number	A29026C
Solution Expiration	210430
Script Version	905
Package Version	IC.5.12.1.RC.1
Start Time	July 21, 2020, 11 a.m.
Completion Time	July 21, 2020, 5:55 p.m.

Chef Template Prep Information:

Chef Last Updated	July 23, 2020, 11:06 a.m.
Chef Instrument Name	CHEF00830
Chef Operation Mode	Customer Mode
Sample Position	1
Tip Rack Barcode	492270059
Chip Type 1	318v2
Chip Type 2	318v2
Chip Expiration 1	None
Chip Expiration 2	None
Templating Kit Type	Ion PGM Hi-Q View Chef Kit
Chef Flexible Workflow	
Reagent Expiration	201231
Reagent Lot Number	2114499
Reagent Part Number	A29901C
Reagent Cartridge Serial Number	None
Solution Lot Number	2112248
Solution Part Number	A25956C
Templating Protocol Planned	Chef Protocol - 200 bp
Solution Cartridge Serial Number	None
Solution Expiration	201031
Templating Protocol Executed	Chef Protocol - 200 bp
Chef Script Version	905
Chef Package Version	IC.5.12.1.RC.1
Start Time	July 22, 2020, 5:02 p.m.
Completion Time	July 23, 2020, 11:06 a.m.

Software Version

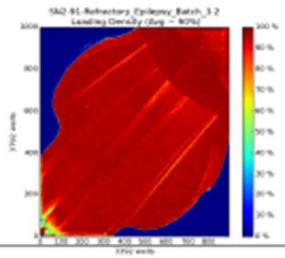
Torrent_Suite	5.12.1
host	5QYTV12
ion-analysis	5.12.27-1
ion-dbreports	5.12.60-1
ion-gpu	5.12.1-1
ion-pipeline	5.12.17-1
ion-torrentpy	5.12.21-1
ion-torrentr	5.12.23-1
Script	22.0.0
LiveView	755
DataCollect	489
OS	21
Graphics	36
Ion_Chef	IC.5.12.1.RC.1

Run Summary

1.24 G	75
Total Bases	Key Signal

90%
ISP Loading

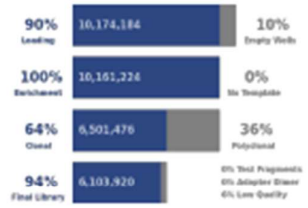
ISP Density



6,103,920
Total Reads

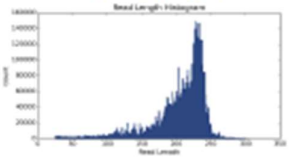
60%
Usable Reads

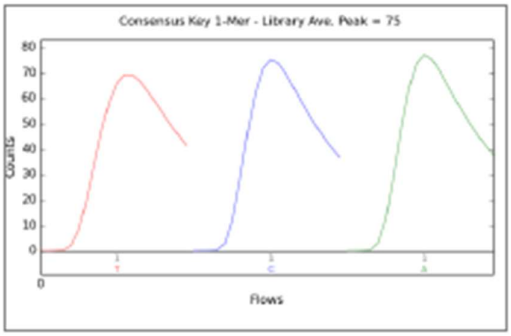
ISP Summary



203 bp	216 bp	228 bp
Mean	Median	Mode

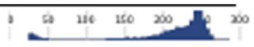
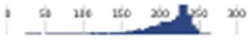
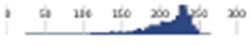
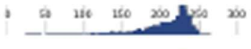
Read Length





Addressable Wells	11,287,275	
With ISPs	10,174,184	90.1%
Live	10,161,224	99.9%
Test Fragment	13,503	00.1%
Library	10,147,721	99.9%

Library ISPs	10,147,721	
Filtered: Polyclonal	3,659,748	36.1%
Filtered: Low Quality	371,103	03.7%
Filtered: Adapter Dimer	12,950	00.1%
Final Library ISPs	**6,103,920**	**60.2%**

Barcode Name	Sample	Bases	≥ Q20	Reads	Mean Read Length	Read Length Histogram
No barcode	none	48,274,935	45,294,910	235,508	204 bp	
IonCode_0125	EE10.1JAD	188,793,431	178,516,545	927,508	203 bp	
IonCode_0126	EE26.1ROB	194,594,208	184,106,455	964,936	201 bp	
IonCode_0127	EE46.1ASH	203,411,141	192,995,784	1,002,143	202 bp	
IonCode_0128	EE95.1JUN	145,329,010	138,049,297	712,816	203 bp	
IonCode_0129	EE100.1PHY	76,532,239	72,345,926	382,928	199 bp	
IonCode_0130	EE128.1SIP	149,667,791	141,965,643	739,493	202 bp	
IonCode_0131	EE138.1CHL	93,751,478	88,817,606	469,661	199 bp	

ion torrent
by Thermo Fisher Scientific

1

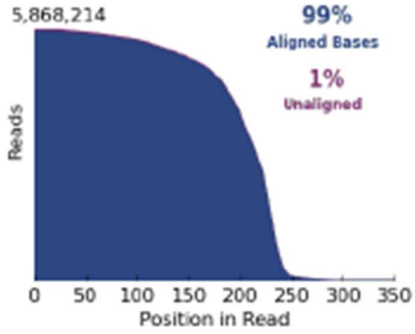
IonCode_0132	UU210.1GAB	135,915,438	128,528,853	668,729	203 bp	

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_1	**11,564**	**91**	

Alignment Summary *(aligned to Homo sapiens)*

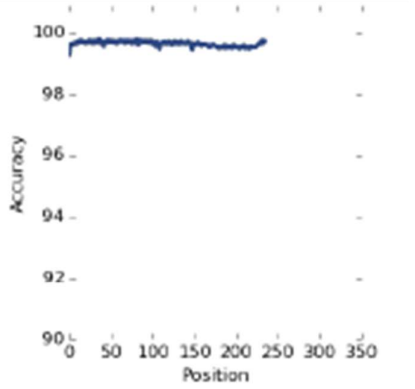
1.18 G
Total Alignment Bases

0.4X
Average Coverage Depth of Reference



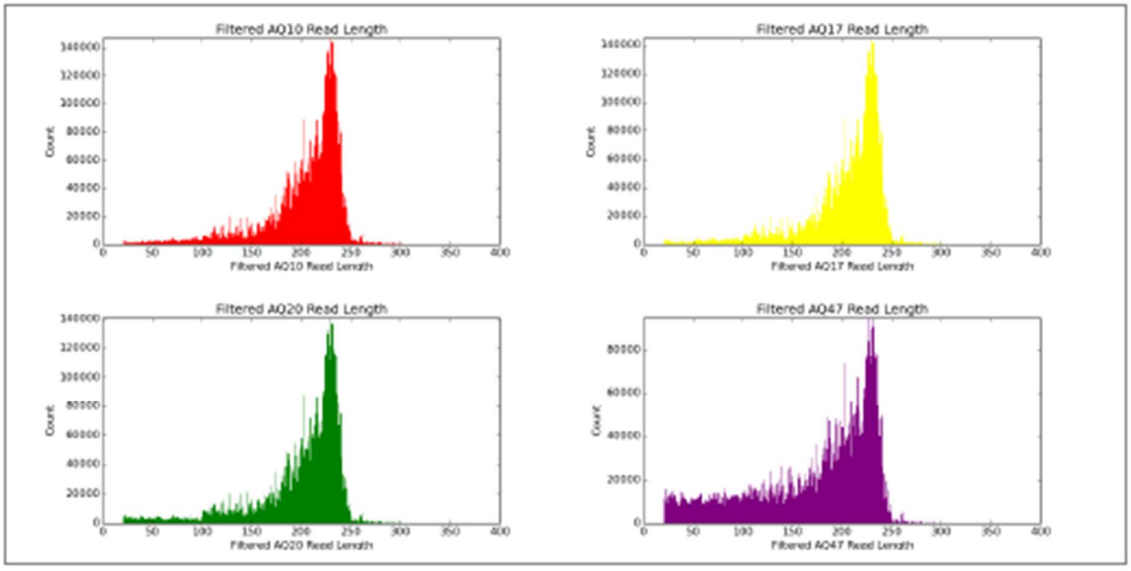
	Count	%
Total Reads	5,868,214	–
Aligned Reads	5,828,509	99.3%
Unaligned Reads	39,705	0.7%

99.7%
Mean Raw Accuracy 1x



Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.16 G	1.13 G	957 M
Mean Length [bp]	201	198	172
Longest Alignment [bp]	339	339	332
Mean Coverage Depth	0.4	0.4	0.3


Filtered AQ10 Read Length
Count
Filtered AQ10 Read Length
Filtered AQ17 Read Length
Count
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Count
Filtered AQ20 Read Length
Filtered AQ47 Read Length
Count
Filtered AQ47 Read Length

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD190273_182_RefracEpilepsySpikeIn

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonCode_0125	EE10.1JAD	921,664	97.02%	449	94.68%
IonCode_0126	EE26.1RCB	956,248	96.96%	463	95.13%
IonCode_0127	EE46.1ASH	996,738	97.62%	486.9	95.06%
IonCode_0128	EE95.1JUN	708,348	97.22%	348.1	94.12%
IonCode_0129	EE100.1PHY	378,451	96.21%	180.5	94.60%
IonCode_0130	EE128.1SIP	730,506	97.57%	356.3	94.71%
IonCode_0131	EE136.1CHL	466,030	97.10%	224.3	92.53%
IonCode_0132	UU210.1GAB	666,524	97.27%	324.7	94.67%

1 5 items per page 1 - 8 of 8 items

ion torrent by Thermo Fisher Scientific

Analysis Details

Run Name	R_2015_02_19_23_39_08_user_SN2-91-Refractory_Epilepsy_Batch_3.2
Run Date	Feb. 19, 2015, 11:39 p.m.
Run Flows	500
Projects	
Samples	EE138.1CHL, EE26.1ROB, EE128.1SIP, EE10.1JAD, EE95.1JUN, UU210.1GAB, EE46.1ASH, EE100.1PHY
Reference	
Instrument	sn274670980
Operation Mode	Customer mode
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Barcode	AB0357404
Chip Check	Passed
Chip Type	318C
Chip Data	single
Chip Lot Number	Missing
Chip Wafer	Missing
Barcode Set	IonCode Barcodes 1-32
Analysis Name	Auto_user_SN2-91-Refractory_Epilepsy_Batch_3.2_291
Analysis Date	July 23, 2020, 9:50 p.m.
Analysis Flows	500
runID	VWRGG
BeadFind Args	justBeadFind –args-json /opt/ion/config/args_318_beadfind.json
Analysis Args	Analysis –args-json /opt/ion/config/args_318_analysis.json –gopt /opt/ion/config/gopt_318v2_Hi-Q.param.json
Pre-BaseCaller	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20
Calibration Args	Calibration
BaseCaller Args	BaseCaller –trim-qual-cutoff 15 –barcode-filter-minreads 20 –phred-table-file /opt/ion/config/phredTable.318.B5.h5
Alignment Args	tmap mapall ... stage1 map4
IonStats Args	ionstats alignment
Analysis Parameters	default

Chef Summary

Chef Library Prep Information for Sample Set Epilepsy Batch 3.2:

Library Prep Type	AmpliSeq on Chef
Library Prep Plate Type	BC 25-32 (Blue)
PCR Plate Serial Number	A040020907
Combined Library Tube Label	C0668518
Last Updated	July 22, 2020, 1:52 a.m.
Instrument Name	CHEF00830
Operation Mode	Customer Mode
Tip Rack Barcode	69924013A
Kit Type	Ion AmpliSeq Kit for Chef DL8
Reagent Lot Number	2189459
Reagent Part Number	A29025C
Reagent Expiration	210331
Solution Lot Number	2144926
Solution Part Number	A29026C
Solution Expiration	210430
Script Version	905
Package Version	IC.5.12.1.RC.1
Start Time	July 21, 2020, 6:56 p.m.
Completion Time	July 22, 2020, 1:52 a.m.

Chef Template Prep Information:

Chef Last Updated	July 23, 2020, 11:06 a.m.
Chef Instrument Name	CHEF00830
Chef Operation Mode	Customer Mode
Sample Position	2
Tip Rack Barcode	492270059
Chip Type 1	318v2
Chip Type 2	318v2
Chip Expiration 1	None
Chip Expiration 2	None
Templating Kit Type	Ion PGM Hi-Q View Chef Kit
Chef Flexible Workflow	
Reagent Expiration	201231
Reagent Lot Number	2114499
Reagent Part Number	A29901C
Reagent Cartridge Serial Number	None
Solution Lot Number	2112248
Solution Part Number	A25956C
Templating Protocol Planned	Chef Protocol - 200 bp
Solution Cartridge Serial Number	None
Solution Expiration	201031
Templating Protocol Executed	Chef Protocol - 200 bp
Chef Script Version	905
Chef Package Version	IC.5.12.1.RC.1
Start Time	July 22, 2020, 5:02 p.m.
Completion Time	July 23, 2020, 11:06 a.m.

Software Version

Torrent_Suite	5.12.1
host	5QYTV12
ion-analysis	5.12.27-1
ion-dbreports	5.12.60-1
ion-gpu	5.12.1-1
ion-pipeline	5.12.17-1
ion-torrentpy	5.12.21-1
ion-torrentr	5.12.23-1
Script	22.0.0
LiveView	755
DataCollect	489
OS	21
Graphics	36
Ion_Chef	IC.5.12.1.RC.1

Appendix H – Condensed Ion Reporter table and *in silico* indicators of pathogenicity for the five verified variants of interest

Gene	NCBI RefSeq	hg19 Genomic Coordinate	Coding Change	Protein Change	Function	ClinVar Designation	rs ID	OMIM ID	SIFT Score	PolyPhen Score	Grantham Score	FATHMM Score	phyloP Score
SCN1A	NM_006920.5	chr2:166852628	c.4444-1C>T	p.?	splice acceptor	Pathogenic	rs1553521567	182389	-	-	-	-	9.81
SCN1A	NM_006920.5	chr2:166909392	c.664C>T	p.Arg222Ter	nonsense	Pathogenic	rs121918624	182389	-	-	-	-	7.76
GRIN2A	NM_001134407.2	chr16:9892299	c.2191G>A	p.Asp731Asn	missense	Pathogenic/Likely pathogenic	rs796052549	138253	0.02	1.0	23.0	0.92	7.68
GABRG2	NM_198903.2	chr5:161520937	c.211A>G	p.Asn71Asp	missense	Uncertain significance	-	137164	0.04	0.145	23.0	-	7.05
GRIN2B	NM_000834.4	chr12:13716673	c.3499G>A	p.Val1167Ile	missense	Uncertain significance	rs1042339	138252	0.4	0.0	29.0	0.94	1.89