Preventive genomics

Bricca, Annalise Rafaela

UNIVERSITY OF ZAGREB
SCHOOL OF MEDICINE

Annalise R. Bricca

Preventive Genomics

GRADUATE THESIS

Zagreb 2021

This graduate thesis was made at the Department of Pediatrics, KBC Zagreb- University Hospital Centre Zagreb and mentored by Prof.dr.sc. Mario Ćuk, MD, PhD. It was submitted for evaluation during the academic year 2020/2021.

Abbreviations

HGP: Human Genome Project

GWAS: Genome-wide association study

SNP: Single nucleotide polymorphism

NGS: Next- generation sequencing

MAG: Medically actionable gene

WES: Whole exome sequencing

CMA: Chromosomal microarray analysis

WGS: Whole genome sequencing

SNV: Single nucleotide variations

INDEL: Insertion and deletion of basis in the genome

MC: Mendelian condition

VUS: Variant of Uncertain Significance

MR: Mendelian randomization

PRS: Polygenic risk score

NIPT: Non-invasive prenatal testing

cffDNA: Cell-free fetal DNA

NBS: Newborn screening

rWGS: Rapid whole genome sequencing

HD: Huntington's disease

mHTT: Mutant huntingtin protein

DMD: Duchenne Muscular Dystrophy

CK: Creatine Kinase

AAV: Adeno-associated virus

Table of Contents

Abstract

Preventive Genomics

Annalise R. Bricca

Preventive genomics may very well be the most advanced emerging field in pediatric medicine to date. Considered to be in its infancy at about 20 years old, the use of genomic medicine became a reality with the completion of the Human Genome Project in 2001. It is a new, developing field and with the potential to predict a child's future health so profound it must be understood as the way to prevent future suffering for children. The birth of a phenotypically healthy baby belies their genetic undercurrent, which may contain mutations manifesting as devastating disorders. Historically, standardized newborn screening has been done on all infants at birth, with the goal of detecting and diagnosing disorders as early as possible. Commonly screened disorders include phenylketonuria, sickle cell disease and cystic fibrosis among other endocrine and metabolic disorders. This screening protocol only focuses on selected disorders and therefore has become limited in scope for early diagnoses of many serious genetic diseases. Traditionally, genome sequencing is only performed after a child has presented with symptoms of a disorder. Preventive genomics has changed this narrative to preemptively detect disorders before any symptom is present and before any indication of illness is found. The medical uses of genomic sequencing in children are myriad but we also found that the ethics of presymptomatic detection of adult-onset disorders has faced controversy. In question are ethical implications of testing when children are not capable of consent, and the long-term effects it may have on them with the awareness of what their medical future holds. In this review, we examine the history of the field of genomic medicine and its potential to change the course of disease, leading to better outcomes in pediatric patients. In particular, we explore how genomics can be used to advance the diagnosis of genetic disorders and decrease the time to diagnosis in utero, infancy, and childhood. We also probe the role of genomic sequencing and its capability to identify disorders far earlier than ever before, allowing for the possibility of gene therapy and cures for those children who are affected.

Sažetak

Preventivna Genomika

Annalise R. Bricca

Preventivna je genomika vjerojatno najnaprednije nastajuće polje dječje medicine do danas. Iako se smatra da je, sa svojih dvadesetak godina, genomika još uvijek u povojima, postala je stvarnost završetkom *Projekta humanog genoma* 2001. godine. Genomika je novo polje medicine u razvoju te ima tako značajan potencijal pri predviđanju zdravlja djeteta nakon rođenja da mora biti prihvaćeno kao način sprečavanja budućih patnji djece. Rođenje fenotipski zdravog djeteta proturječi genetskom podrijetlu ako sadržava mutacije koje se manifestiraju kao razorni poremećaji. Povijesno gledano, novorođenčad je podvrgnuta standardiziranom probiru pri rođenju kako bi se mogući poremećaji što ranije otkrili i dijagnosticirali. Ovakav probir novorođenčadi, uz ostale endokrine i metaboličke poremećaje, obično uključuje fenilketonuriju, bolest srpastih stanica i cističnu fibrozu. Budući da ovaj protokol probira u središtu pozornosti ima samo odabrane poremećaje, postao je ograničen u ranoj dijagnostici mnogih ozbiljnih genetskih bolesti. Tradicionalno, sekvencioniranje genoma provodi se tek nakon što dijete pokaže simptome poremećaja. Preventivna genomika promijenila je ovakav pristup dijagnostici jer se poremećaj otkriva preventivno, prije nego što je prisutan bilo koji simptom ili naznaka bolesti. Bezbroj je medicinskih uporaba genomskog sekvencioniranja kod djece, međutim presimptomatsko otkrivanje poremećaja izazvalo je kontroverze kod odraslih. U pitanju su etičke posljedice testiranja, s obzirom na to da djeca nisu sposobna pristati na testiranje te na dugoročne učinke koje to može imati na njih ako je unaprijed jasno kakva je njihova medicinska budućnost. Ovaj pregledni rad predstavit će povijest genomske medicine i njezin potencijal mijenjanja tijeka bolesti što dovodi do boljih ishoda u liječenju djece. Naglasak će biti na istraživanju korištenja genomike za napredovanje u dijagnostici genetskih poremećaja te smanjenju vremena za dijagnostiku u trudnoći, dojenačkoj dobi i djetinjstvu. Također, ispitat će se uloga genetskog sekvencioniranja i njegova sposobnost prepoznavanja poremećaja mnogo ranije no ikada prije što dopušta mogućnost genske terapije i liječenja pogođene djece.

Introduction and History of Preventive Genomics

Preventive genomics in medicine would frankly be non-existent if it weren't for the Human Genome Project (HGP). Beginning in 1990 and spanning 13 years, this international collaboration yielded the first ever DNA sequence of the human genome, comprising three billion nucleotides. A person's genome makes up all of their DNA, comprising four nucleotide bases (Adenine, Guanine, Cytosine and Thymine) and genes that contain instructions for making proteins in the body. Our DNA gets wrapped around histone proteins and makes chromosomes, which lie within the nucleus of our cells. Humans have 23 pairs of chromosomes in each cell, and a total of 46 chromosomes. The significance of this explanation is to illustrate that before the human genome was completely sequenced, there was the concept of individual genes, separate from each other, that coded for proteins and the rest of the DNA was known as "junk" DNA. With sequencing, we came to understand that greater than 80% of the human genome is functional and a considerable amount involved in gene expression (1).

The moniker "genome" is in fact attributed to a German botanist by the name of Hans Winkler in 1920, as he made a hybrid term to fuse words "gene" with "chromosome". Lederbert and McCray later disputed this origin however, and in 2001 claimed that adding the suffix "-ome" to any word (in this case to the word "gene") makes reference to all genes of an organism. It very well may be that both are equally right; by the 20th century the concept of chromosomes containing all of our genes was already well-established (2). However, it wasn't until the inception of the Human Genome Project in the 1990s that we had one completely sequenced human genome which could act as a "reference" sequence for all human DNA, as we know that any two random individuals share approximately 99.5% of their genome sequences (3).

The sequenced human genome has lead to genome-wide association studies (GWAS) looking for single nucleotide polymorphisms (SNPs) within the genome in order to study many people, some

who have and some who do not have a common condition (3, 4). SNPs are the change in a single nucleotide within the DNA sequence, which changes one base pair. These are the most often seen genetic variations in people, usually not harmful. These variants (formerly commonly referred to as "mutations") are studied to see whether or not there is a disease associated with such genetic variation. In 2005, a herald GWAS was published and showed a relationship between age-related macular degeneration and variants in the complement factor H gene. This genetic variant-disease relationship was not known until the GWAS was published, and it became the harbinger of more and more large studies looking at conditions that are relatively common in pediatric populations, including Type 1 Diabetes and asthma (3). In the 15 years since this first revolutionary study, many more associations have been established between gene mutations and neurological or developmental disorders. Over 20,000 human exomes have been sequenced yielding the discovery of hundreds of genes linked to rare, single-gene Mendelian diseases (1). The use of next-generation sequencing (NGS), also known as massively parallel sequencing or high throughput sequencing, became clinically applicable in 2005. NGS refers to the "highly parallel or high-output sequencing methods that produce data at or beyond the genome scale" (4). This means NGS is sequencing the protein-coding part of the genome, known as the exome, or sequencing the whole genome including protein-coding and non-coding parts. This new and advanced technique competed, in a sense, with microarrays that had been used since 2003. Microarrays looked for whether an individual's DNA would bind to a mutated DNA sample in order to confirm present mutations (5). Chromosomal microarrays are still used generally as a first-line test for congenital anomalies or known syndromes, such as 22q11.2 deletion syndrome (6). The advent of microarrays and next-generation sequencing told scientists not only more about individual genomes but in particular about disease-associated genetic variations that had not been exposed before (5).

Only a mere 50 years earlier had scientists James Watson and Francis Crick presented the DNA double helix to the public for the first time, introducing the deoxyribonucleic acid structure containing genes in every human (8). The significance of knowing the sequence of nucleotides in a genome it that it provides the genetic information that is unique to each person; that 0.5% of

the genome sequence not shared by any two people is what makes us like no one else on the planet (8). This tiny percentage of genome sequence variation can contain some changes. These can be benign variations such as SNPs; usually these point mutations don't affect how the genome functions or its structure since they tend to occur in noncoding regions of DNA (9) . These occur in about 1 in every 1,000 bases of the genetic code. Conversely, entire chromosomal deletion results in disorders such as Prader-Willi, or adding of a chromosome yields Trisomy 21 (1). The percentage of variation in an individual's genome in fact may determine their lot in life; whether or not they face a future of developmental difficulties, phenotypic differences or pain and suffering. Th role of preventive genomics in these genetic variations is to identify conditions, whether primary or secondary, before children manifest phenotypic differences in development and become symptomatic (2).

A common theme we found was that genomic medicine is now more than ever breaking through its previously specific and few rare clinical indications, and "poised to go mainstream" (5). The assertion by one author that "genomics is a scientifically based fortune-teller" has not been so far from the truth (10). To diagnose conditions and prevent complications before they ever occur has the potential to ameliorate quality of life for prenatal, infant and pediatric patients and is groundbreaking for children with genetic variants destined for certain disorders (11). Though the field is new and developing, its promise to "see" into the future is a hopeful prospect for many families whose children are disabled or yet undiagnosed. Identifying the susceptibility of a child to certain diseases when they are asymptomatic and before a family embarks of an oft experienced "diagnostic odyssey" is one of the most tangible goals of preventive genomics (5). Modern-day genomics also opens the door to offering precision or "personalized" medicine to patients via tailored genetic therapies for gene variants in "medically-actionable genes" (MAGs) found in single-gene Mendelian diseases, or disease-causing monogenic or syndromic variants (12). The development of this field has granted quicker diagnoses, earlier and more effective treatment and less agony for pediatric patients in ways that could not have been imagined before two decades time.

Next Generation Sequencing

Next generation sequencing (NGS) also known as "high throughput data" has transformed how we detect genetic variations in a person and their susceptibility or risk of disease. However, sequencing had its roots much earlier than when NGS broke out onto the scene. The first sequencing used was 1977 was Sanger sequencing. Although able to detect single-gene (Mendelian) disorders quite well such as hemochromatosis, sickle cell anemia and cystic fibrosis, it didn't detect structural rearrangements, duplications or deletions in the genome (13). Sanger sequencing was used widely for a long period of time. Now since its clinical induction in 2005, the use of NGS has driven genomic discoveries by its ability to detect very small variants in a genome, which the previously popular DNA microarrays and Sanger sequencing could not do (5). The popularity and widespread use of NGS in the past 15 years can be attributed to its far-reaching recognition of structural and sequence variants in the genome, which could be translocations, duplications, inversions or deletions (13). Why NGS has become so groundbreaking for genomics, though, is its capability to run "massively parallel sequencing," which processes millions of DNA sequences in parallel, or simultaneously (14). This means much faster sequencing, faster results and is cheaper as well, earning NGS the colloquialism as "the $1000 genome" (5).

Exome Sequencing

Exome sequencing or whole exome sequencing (WES) refers to analyzing the protein coding regions of the genome; that is, all of the exon regions put together which remarkably contain approximately 85% of disease-causing variants (14). The exome is only about 1% of the entire genome but just sequencing these areas has a high yield of information, both for identifying new variants or identifying variants known to cause disease. Exome sequencing is known for identifying certain disorders; for example, one article cited a 29-55% diagnostic yield for neurodevelopment disorders (14). Generally, the diagnostic yield is anywhere between 10-50%. WES information integrated with a phenotype, or trait, that we can characterize in a person's appearance or behavior is another way that sequencing can be used not only to provide

information about the genome but how that affects the clinically presentation of the patient (15). More succinctly, knowing a phenotype and having precise genomic information can help to narrow the breadth of diagnoses to those associated with particular genes and diseases (16). An illustrative example is a 2019 study from Germany that employed exome sequencing and clinical phenotyping in 50 children with developmental abnormalities or undiagnosed neurological illnesses. In 42% of the children there was a gene mutation associated with a disease, leading to disease identification. Not all variants lead to a disease we can pinpoint, however, and 44% of the children had changes in genes not known to be associated with disease (17). The utility of this example is to see that still close to half of the undiagnosed children with disorders did receive a diagnosis after WES. This outcome is corroborated by a recent metaanalysis in 2019 by Srivastava et al (18), where they reviewed 30 articles that had data on the diagnostic yield of WES for neurodevelopment disorders. In summation, WES was diagnostic in 36% overall . This data led to the recommendation WES should be a first-line test for the group of neurodevelopment disorders, whereas previously chromosomal microarrays (CMA) had been the standard first -line test with a much lower diagnostic yield of 15-20% (18). WES has shown an undeniably higher yield for diagnosing children in these cases, firmly earning its place in NGS.

Whole Genome Sequencing

The most comprehensive of the NGS techniques, whole genome sequencing (WGS) does exactly as its name suggests: it sequences the entire genome, not only targeting protein coding regions. Its used to detect a large number of mutations including balanced translocations, inversions and short tandem repeats, like those seen in Huntington's disease. One drawback which was repeated across sources was the fact that WGS, while being the most all-encompassing of the NGS methods, has "limited standardization" or limited "comparative data" through which to make it useable clinically. This ultimately has led to WGS being used mainly for detecting single nucleotide variations (SNVs) and small insertions and deletions of bases in the genome (INDELs) (19, 20). That being said, WGS has also produced results that indicate it may reveal more about a person's genome than WES alone. In a Canadian study of 103 pediatric patients

with phenotypes suggestive of genetic disorder, WGS analysis was compared with their previous genetic testing results, including CMA and WES results. WGS detected diagnostic variants in 42 of the participants, significantly more than had a molecular diagnosis from conventional genetic testing, of which only 25 participants had. WGS not only found all the variants and CNVs from the patients' previous WES, in addition 17 patients had diagnoses that were only made by WGS. Further, WGS found diagnostic variants not identified by the WES in patients who were diagnosed by WGS, a total of 25% of the diagnosed patients. (20) This data is compelling evidence that WGS has a higher diagnostic yield than other NGS methods and is especially valuable in the testing of children in which we suspect a genetic disorder. The results suggest that WGS has relevance from the outset of testing in a child with a suspected genetic disorder or phenotype, making the process much less drawn out (20). This makes the case for WGS to be used in clinical practice and not relegated to such narrow indications for use.

Targeted Sequencing

Targeted sequencing or gene panel sequencing is a type of genomic sequencing that uses a gene panel made up of only specific genes to be sequenced, not the whole genome. Generally these genes have a strong correlation with a disease, and targeted sequencing delves into much greater sequencing depth (21). Due to their specific results, gene panels are commonly used in trials or in research to create targeted therapies. They generate less data than WGS and WES but are very rapid and similarly show genetic variants, both common and rare, for the selected genes (22).

The figure (*Figure 1*) below shows four statistical representations of gene discoveries made over a period of time. *Figure 1*A shows the approximate rates of gene discoveries for Mendelian conditions (MCs), showing especially high peaks from 2010-2020 with the advent of exome sequencing. *Figure 1*B illustrates the number of gene discoveries by method, either conventional or next-generation sequencing, from 1996-2019. *Figure 1*C delineates the gene discoveries per

year for MCs by their mode of inheritance, and also *de novo* variants that were discovered, between 1996-2019. *Figure 1*D compares the number of new phenotypes reported per year during 1996-2019 between classical syndrome delineation (in orange; phenotype-driven) versus delineation based off of genotype-driven syndromes (in teal), where people are identified to have a syndrome only after finding out they have a pathogenic variant of the same gene (23).

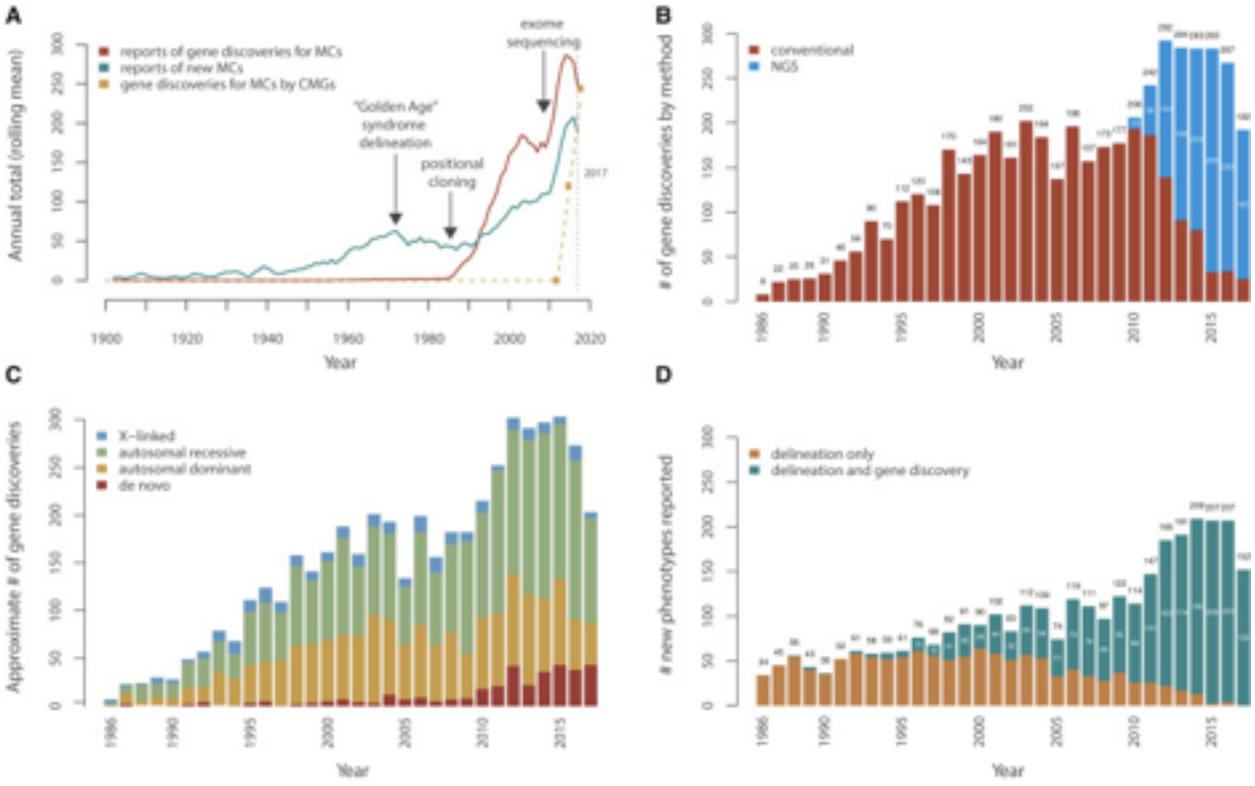

Figure 1. Annualized Metrics of Gene Discovery for Mendelian Conditions. According to: Bamshad et al. (2019), p. 448-455 (23).

Interpretation of next generation sequencing data (NGS)

The primary challenge once data is gathered is determining how significant the identified variants are in predicting disease. NGS produces so many SNPs and copy number variants that it can be difficult to decipher significance to all variants . The sequence is compared against the reference genome to detect variants and only retain "high quality variants", rare variants, and those which are expected to impact a gene's protein coding sequence such as frameshift, nonsense, and missense mutations. False positives can occur due to mistakes in sequencing and usually are easily detected and filtered out. There is not always agreement in laboratories on how to designate pathogenicity of variants; the can be determined to be either pathogenic, likely pathogenic, a variant of uncertain significance (VUS), likely benign or benign. This can make interpretation difficult especially if a variant is considered "potentially pathogenic" but it hasn't been determined whether or not it has any association with a genetic syndrome (18). Several sources made reference to using transcriptome assays as a complementary analysis to genome sequencing (11, 14, 24). Transcriptome analysis involves using RNA transcripts that are made by the genome as a kind of sign of gene activity in a cell. Transcriptome data has also improved diagnostic yield when used in conjunction with sequencing; in a study with 50 people with undiagnosed muscle diseases, using RNA sequencing resulted in 17 new diagnoses after DNA sequencing gave no genomic information to assist in diagnosis (25).

Analysis of an individual's genome with the use of WES or WGS generates large amounts of data that can be used to associate mutations with diseases, diagnose a child with a rare disease or tell us their risk of developing a disease. Still, the sheer volume of the data can be difficult to interpret. There may be many variants when compared to a reference genome, but the sequencing also gives more VUS', which leave uncertainty and room for different interpretation (2, 26).

Some VUS's are being reexamined with genomic testing after a period of several years to see if the VUS status changes or remains the same, using the updated standards for genetic variant

testing (27). In a 2017 retrospective study from Texas, NGS was performed to re-test children who had been diagnosed with epilepsy from 2012-2015 to see if their variant status changed from its initial classification. Clinically significant results were defined as a Pathogenic-Likely Pathogenic (P-LP) variant which had changed to a VUS or Benign-Likely Benign (B-LB) status, or a VUS that was upgraded to P-LP status. Out of 185 patients tested with a genetic variant and mean age of 5 1/2 years old, the results showed 36.2% of patients had a gene variant reclassified. Significantly, 19 P-LP variants became classified as less pathogenic (14 VUS and 5 B-lB) and 2 patients had their variant pathogenicity increased from VUS to P-LP. Out of the 124 patients who did not have a diagnosis and were previously classified as VUS, only 1.6% were upgraded to P-LP but interestingly 37.1% were downgraded to B-LB. This study demonstrates how dynamic genomic testing can be, and is the basis for the study's recommendation to re-examine patients with variant status every 2 years for the possibility of re-classification (27). It also illustrates genome plasticity; whether influenced by environment or other factors, the genome has the ability to change over time. Thus, reclassification of variants can influence therapies and at the very least gives a patient a current idea of what their variant status classification is.

DNA Microarrays

With NGS breaking out onto the genomic testing scene in 2005, DNA microarrays remained in steady use although without the fanfare of NGS. Also referred to as "gene chips," DNA microarrays are a fragmented piece of DNA that is put on a chip platform with a known DNA sequence. If the individual's DNA will fail to bind with the synthetic DNA on the chip then it indicates the person has a mutation. DNA microarrays were widely used in the 1980s and 1990s, and clinical microarray testing began in 2003 (1, 5). These tests are reliable especially for finding single nucleotide changes at certain locations but are limited in scope since they are not necessarily capable of detecting genome variations (5). DNA microarrays have long been a mainstay for their high accuracy and low cost. They now are replacing karyotyping as the first-line testing for children with congenital anomalies and developmental abnormalities, as they are

able to detect chromosomal structural changes (5, 28). Articles that we found elaborated far more about NGS technology and uses, and DNA microarrays were often not mentioned. Overall, they are seemingly an older but effective technology though far less comprehensive compared to the current NGS.

Genome-Wide Association Studies

Genome-wide association studies (GWASs) since their inception in 2005 have utilized large cohorts of people to use as test subjects and are responsible for creating a massive library of genetic variants (in the millions). These variants include many that have associations with particular diseases or traits. One of the most significant parts of these associations is they contain no "a priori" hypothesis; that is, there is no subjectivity and the data is collected independently. This has swung open the door to discover new gene and disease associations that have not been noted before, and a tremendous increase in gene discovery (29). As of 2020, 4500 GWASs have been reported upon and have discovered more than 55,000 loci for about 5000 traits and diseases (29, 30). We found that GWAS' were favorably regarded, however, criticisms included that only a very small amount of the SNP-trait or disease associations have been further probed and therefore not as useful without knowing "causal variants, target genes, and the underlying mechanisms linking the variants and genes to the original phenotype" (31). Other defects of GWAS included a lack of non-European participants; the overwhelming majority in GWAS are of European ancestry with only 10% of other origin. Limiting results to only a certain population of people results in limited genetic research for others and thus can transfer over into a lower quality of clinical care, according to critics (29).

The GWAS exists to also help us look at causality, exposure and lifetime risk of developing a disease. Mendelian Randomization (MR) and Polygenic Risk Scores (PRSs) are two of the most important ways that the GWAS' come to be used clinically. MR is used on an already established relationship between a risk factor and disease. The risk factor could be LDL cholesterol and triglycerides and evaluating their causality with coronary heart disease, for example. The goal of MR is to clarify whether these are relevant associations or to rebut them. (30).

Polygenic risk scores (PRSs) evaluate causality between an exposure and development of a disease. Genetic variants that strongly are associated with exposure and disease would result in an "inferred" causality between the two. PRSs are another use of GWAS'; they calculate an individual's genetic risk of developing disease in their lifetime. PRSs use the compiled information from GWAS' to give a summary of all the effects of a large number of genetic variants and their effect on increasing an individual's risk of disease (32, 33). Many sources stressed that these "risk scores" must not be used in a vacuum, rather they give us information that could be useful or predictive but should be "assessed in the context of existing clinical predictors of risk," helping clinicians and patients to see what lifestyle modifications might need to be taken to help with prevention (29).

Preventive Genomics : Prenatal and Infancy

Next Generation Sequencing in utero and infancy

Prenatal screening tests began in the 1970s and now standard fetal testing is called NIPT, or non-invasive prenatal testing. This is testing for fetal chromosomal abnormalities or chromosomal aneuploidies, including Trisomy 21, 18 and 13. NIPT is considered an "unconfirmed diagnostic test" and the research done thus far is inconclusive as to whether it can detect entire genome regions or sex chromosome aneuploidies. To be positive, a chorionic villus sampling or amniocentesis in the second-trimester is required to confirm the fetal karyotype (34). The launch of NGS for fetal whole-genome sequencing in 2015 began after Lo et al. identified the cell free fetal DNA (cffDNA) circulating freely in maternal plasma, noteworthy because it is very highly fragmented with each fragment containing about 50-200 base pairs (34, 35). Studies have shown that these shorter fragments have more cffDNA and provide a decreased probability of a "no-calls" result, where if greater than 5% of SNPs are missing, the result is not deemed significant (26). This new cffDNA testing obscured invasive techniques such as amniocentesis and chorionic villus sampling that had been used to confirm chromosomal abnormalities. Boasting a 95% sensitivity and specificity rate, cffDNA analysis is a powerful yet non invasive method for prenatal diagnosis (34).

Prenatal imaging for discerning phenotypic abnormalities in utero is standard practice followed by chromosomal microarray analysis (CMA) and karyotyping to identify anomalies such as aneuploidy, duplications or deletions. Approximately 30-40% of those tested have the anomaly identified, which leaves 60% undiagnosed. In these cases, prenatal genomic sequencing can increase diagnostic yield significantly; a study by Normand et al. showed 35% diagnostic yield from prenatal exome sequencing in previously undiagnosed patients who had already done CMA or karyotyping (11). Such an increase in diagnoses makes a strong case for including genomic sequencing to detect fetal abnormalities, ultimately more rapidly giving important information to the parents and influencing therapy, monitoring and planning.

A retrospective study in 2018 examined fetal exome sequencing from 2012-2017 after fetal ultrasound detected some structural anomaly in which Mendelian etiology was suspected. The overall molecular diagnostic rate was 32% from this prenatal exome analysis, of which 50% were autosomal dominant disorders, 41% were autosomal recessive and 9% X-linked. Among the disorders diagnosed were Osteogenesis Imperfecta types 1-4, Ehlers-Danlos syndrome, Marfan syndrome and Cornelia de Lange syndrome type 1. The highest diagnostic yield was among fetuses whose ultrasound showed anomalies affecting more than one organ system and those with craniofacial abnormalities. The study concluded that the exome sequencing was very useful for fetal diagnosis when there was a structural anomaly present and in addition these diagnoses were obtained after the usual standard prenatal tests did not yield any (36). This points to the utility and necessity of next generation whole genome sequencing for early diagnostics that no other standard prenatal genetic tests are detecting.

The following figures, A and B, correspond to the retrospective study described in the paragraph above. Figure A is a bar graph showing the association between the greater number of affected organ systems leading to higher diagnostic rate. Figure B shows diagnostic rate as a function of organ system involvement. Figure B shows molecular diagnostic rate of fetuses with (+) or without (-) abnormalities in their corresponding organ system. Fetuses with craniofacial abnormalities had a significantly elevated diagnostic rate compared to the other organ systems shown (36).

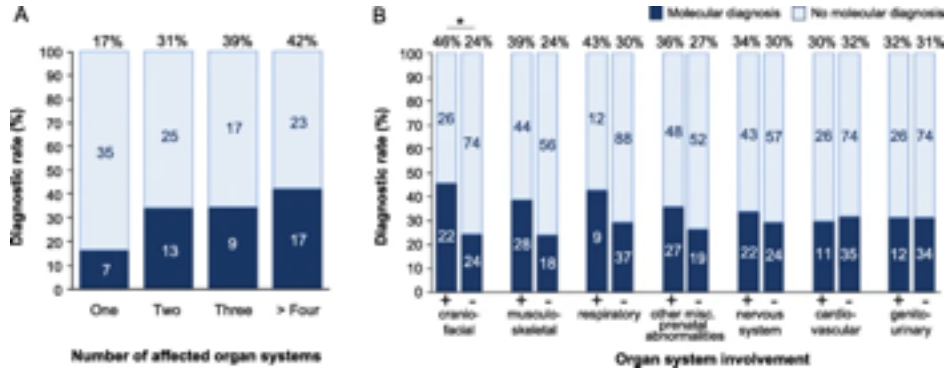

Figure 2. Clinical exome sequencing for fetuses with ultrasound abnormalities and a suspected Mendelian disorder. According to Normand EA, et al. (2018), p. 74.(36).

Despite these diagnostic advances, genome sequencing of fetuses and neonates has not been routine. Patients deemed "high risk" are exclusively indicated for these tests, usually due to advanced age at pregnancy. The American College of Medical Genetics and Genomics (ACMG) has released guidelines for genome sequencing in the infant population, which states the indication for infants is only when they already present with a "likely genetic disorder" phenotypically, a diagnosis, and other testing has not yet identified the exact disorder (37). The data, however, have given momentum to those advocating for whole genome sequencing to detect genetic disorders in utero. A study by Ehrich et al. used whole genome sequencing-based NIPT on more than 10,000 women deemed to have high-risk pregnancies due to advanced age, abnormal ultrasound findings or other indications. The results showed almost double the women tested positive for genetic abnormalities, with 5.4 % testing positive with WGS compared to only 2.3% using other NIPT techniques (38). Consequently, WGS shows higher yield with fetal and neonatal disorder detection than any other currently used screening techniques.

Preventive Newborn Genome Sequencing

Preventive genomics is perhaps no better utilized than at the beginning of life to discover diseases before they materialize in healthy infants and also to find a diagnosis for sick infants (26). Universal newborn screening programs (NBS) with neonatal blood spots are standard practice, yet they only detect a handful of metabolic, endocrine, and hemoglobin disorders (39). The first newborn screening began in 1962 with the heel prick Guthrie test for phenylketonuria. It was the first example of newborn screening that had the potential to save lives and prevent disability through a simple heel prick blood sample (2). Before this, children with deficiency of the enzyme phenylalanine hydroxylase would have a buildup of phenylalanine with devastating consequences. Often not diagnosed until a child presented with a musty odor, hypopigmentation, seizures or mental disabilities, this fate was practically eradicated with the Guthrie test. By the

1990s, tandem mass spectrometry had expanded to almost 30 primary and 26 secondary conditions which have made up the basis of NBS ever since.

Although preventive in nature, NBS tests detect no monogenic conditions and do not help in identifying many rare diseases, leaving many children undiagnosed for years. A rare disease is defined as affecting less than 200,000 people in the United States, according to the Rare Disease Act of 2002. However rare they may be, all together rare diseases affect almost one in ten Americans. And approximately 80% of these have a genetic etiology. Children suffer the brunt of this as they are half of the people with genetic disease. The Global Genes Project had estimated that 30% of them die by the age of 5 (37). The National Institutes of Health reported that children with congenital diseases were most frequently diagnosed from ages 4-6 and teens within ages 16-18, who already presented with symptoms earlier in childhood (16). Symptomatic children at least tell us that something is not normal, but asymptomatic children offer no clues of future disease. Even for babies who have been sequenced, it can be difficult to interpret the clinical picture when a seemingly healthy newborn baby has no symptoms or phenotypic presentation. This is often the situation with diseases with later childhood onset.

A study published in 2016 including almost 1700 neonates and their parents was conducted from 2011-2014 to see the utility of whole-genome sequencing to detect disorders based off genes that were analyzed by NBS. The goal was to see the potential of WGS compared to NBS. WGS was done for the neonate and both parents. All conditions covered by NBS were included; the genes selected and mode of inheritance associated with them as well as associated conditions were chosen from several genomic databases (36). The results confirmed that WGS had less false positives than NBS, and clarified results that had not been conclusive from NBS. Among its strengths is that WGS can identify those who are affected or at risk of developing many more disorders than NBS. Of vital importance is WGS' role in identifying causal mutations for those who were affected. The conclusions from this study agreed with proponents for preventive genomics that sequencing in neonates takes NBS to a majorly expanded level of screening.

WGS shows the broadest range of detection for conditions where early diagnosis can be crucial such as in Mendelian disorders, and also detects a larger quantity of mutations (36).

Proponents of NGS in infants and children advocate that it would expand the breadth and depth of this testing to include Mendelian disorders and reveal more gene and disease associations, ultimately decreasing cost and length of hospitalization for children with "chronic complex conditions," such as those undiagnosed (2, 37) . The years spent living with undiagnosed diseases, in some cases with chronic symptoms and without answers, has fueled preventive genomics to support whole genome sequencing on infants in order to put an end to the drawn-out process. For parents as well, a specific diagnosis in utero or in infancy gives a general prognosis of the patient's future. It can be a relief to know what their condition may bring and be able to plan care and treatment their child. Ultimately preventive genomics is not only detection of disease but knowing earlier and preparing for it. This allows parents and clinicians to give the patient the most opportunities to thrive (14, 18). Importantly, knowing a diagnosis can impact family planning and by parents understanding genetic etiologies they can make more informed decisions about having future children .

Mendelian Diseases

Mendelian diseases are monogenic, or single gene disorders resulting from a mutated gene that is usually inherited. In 1865, Gregor Mendel's modes of biological inheritance were proposed as traits passed down through generations but in different patterns. Regardless of whether the mode of inheritance is autosomal dominant, recessive or X-linked, there are many feared diseases that can strike children early in life. Though over 7,000 Mendelian conditions have been identified, they are considered to be rare. However, these rare conditions are now much better understood. Since the first genome was sequenced and the use of NGS rose, the number of recorded genetic variants has skyrocketed and completely changed clinical diagnoses (40). These were the beginnings of being able to sequence a patient and family's exome or genome and have a definitive diagnosis of a monogenic disease. The great leap from having a rare, undiagnosed

disease to a diagnosed disease again changes the course of one's life, with genomic sequencing offering a glimmer of hope for some improvement with targeted therapeutics (35).

Mendelian diseases also unfortunately contribute to infant deaths in developed countries, up to 23% according to one study, even though they are only present in about 6% of infant births (40,41). These genetic disorders are adding to child mortality and especially when not identified quickly. In the neonatal intensive care unit (NICU), there have been immense efforts to obtain diagnosis of rare genetic disorders within 24 hours with rapid whole genome sequencing (rWGS). The results of these efforts can be life-saving and decrease morbidity and mortality. A retrospective study illustrated the utility of rWGS with a cohort of 42 infants in the NICU classified as being acutely ill. With a diagnostic sensitivity of 42%, the rWGS diagnosed 18 of the 42 infants with 19 genetic diseases. 16 of the infants' diagnoses confirmed the associated phenotype. The short-term outcomes were that 26% of the infants with rWGS diagnoses survived, one's chance of mortality reduced by 43% and one began palliative care (41). The clinical value of early genetic diagnosis for these critically ill infants was "measured by acute precision medicine interventions" that directed the course of treatment. The outcomes were improved by using rWGS and the quick turnaround in providing a genetic diagnosis prevented misguided treatments, thus reducing the length of stay for those who received precision medical interventions.

The table (*Table 1)* below shows acute precision medicine interventions in 13 of the 18 infants receiving genetic disease diagnoses and the resultant changes in outcomes including medication change, change in surgery, palliative care initiated, imaging or procedure change, morbidity avoided and mortality avoided (41).

Table 1. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. According to: Farnaes, L. et al. (2018); (41).

Infant ID	Causal gene	Medication change	Change in surgery	Palliative care initiated	Imaging or procedure change	Morbidity avoided	Mortality avoided
5011	*NPC1*	Miglustat started				Neurologic damage delayed	-
5012	*ARID1B*			Yes		Further futile intensive care	-
5014	*NEB*		Avoided muscle biopsy		Avoided EMG and NCS	Anaesthesia and muscle biopsy	-
5018	*POLR3C*				MRI of brain recommended		-
5019	*GABRA1*	Steroids weaned; confidence in therapy when readmitted			Avoided repeat EEG	Discontinuation of appropriate anti-epileptic at next admission	-
5020	*TPM1*		Cleared for cardiac transplant			Delay in heart transplant	-
5021	*PCDH19*	Ganaxolone started; confidence in medications for child and sibling					-
5024	*PHEX*	Start phosphate and high-dose calcitriol				Development of rickets	-
5026	*JAG1*		Avoided Kasai hepatoportoenterostomy			Kasai and liver transplant	83–94% decrease
5030	*NF1*				Brain MRI for tumour evaluation and MR angiography of renal arteries for stenosis	Potential early detection of NF1 associated tumors	-
5041	*KCNQ2*	Carbamazepine started; phenobarbital weaned				Prolonged uncontrolled seizures with potential neurological damage	-
5053	*ABCC8*		Earlier partial pancreatectomy			3 additional weeks of hypoglycaemia with potential neurological damage	-
5056	*ACTG2*	Started cisapride					
Total		5 (38%)	4 (22%)	1 (6%)	4 (22%)	10 (56%)	1 (6%)

Ethically, there are those who oppose this preventive look into the future of a child's health. There have been questions about when and if genomic sequencing should be done to detect Mendelian disorders with adult-onset symptoms (2). Some argue that it is for the child's best interests to know, while others advocate that knowing they will have future onset of a disease "may hamper children's right to an open future," or "violate their right (not) to know," as well as negatively impact their self-esteem (42).

Huntington's Disease

One monogenic disease usually concealed until adulthood is Huntington's disease (HD), a neurodegenerative disorder affecting a patient's motor, behavioral and cognitive function where death is usually within 15 years of diagnosis. Patients face a devastating prognosis with deterioration of basal ganglia, striatum and cortex. The disease's well-known uncontrollable 'dancing' chorea signals the loss of autonomous control of body and mind simultaneously (43). A disease of autosomal dominant inheritance, HD's genomic hallmark is trinucleotide CAG repeats in exon 1 of *HTT* gene near the N Terminus, which encodes the huntingin protein. This expanded CAG repeat then causes a mutant huntingin protein to be made (*mHTT)*. Children are asymptomatic usually, and age of onset varies as an adult. Research shows that the age of onset of motor symptoms is inversely related to the number of CAG repeats (44). HD displays anticipation, a phenomenon where the *HTT* gene can increase in size with every generation and further elongate the CAG repeats. The cruel impact of anticipation is that with each generation, onset is younger and younger. A study showed that the encoded glutamines which get expanded by CAG's repetition on huntingtin protein were not driving the age of onset, but length of the expanded CAG repeat on *HTT* gene was (43). Research into treatment for HD is majorly focused on decreasing the *mHTT* protein, whose presence is thought to be responsible for the neurotoxic destruction taking place in the brain (45).

Approximately 60% of the variation in age of onset of HD is due to the length of *HTT* CAG repeat while the other 40% of variation is due to heritability. A GWAS was done to determine what factors were disease-modifying before onset and if they accelerated or decelerated onset. In the 4,000 people with HD in the study, modifier alleles were identified as infrequent but with a strong impact and low-impact common modifiers were found acting in new loci. These modifier alleles can influence and change expression of another gene. The conclusion was that genetic modifier loci are evidence of "DNA maintenance mechanisms" being involved in time of onset of HD (43). DNA maintenance mechanisms are responsible for repairing and keeping the genome stable; if they don't work properly it can lead to genomic instability and damage.

Huntington's disease therapeutic genome editing

Given that most children don't have symptoms until they get older, researchers have suggested "DNA maintenance modifier genes" could be having an influence on the CAG repeat's increasing length in somatic cells on *HTT* gene. This could be a possible target of therapy to prevent or slow down the onset of HD (43). To slow down the production of *mHTT*, one approach is to modify gene transcription so as not to transcribe *HTT* into HTTmRNA which would then make the toxic protein. Several modes of targeted genome editing are being researched including CRISPR/Cas. CRISPR (clustered regularly interspaced short palindromic repeats) is a series of DNA sequences found in genomes of bacteria; they are used to detect the segment of DNA that we want to remove. Then, Cas9 protein cuts out the undesirable segment. CRISPR/Cas genome editing is another approach aimed at targeted use of Cas9 protein's RNA-guided nuclease function to excise CAG repeats from the mutated *HTT* and generate two wild-type *HTT* alleles. CRISPR/Cas9 could also function to edit the genome by inactivating the mutant allele, causing a null state (45).

The following figure *(Figure 3)* shows two images depicting brain tissue in a mouse model of Huntington's disease. The upper image shows abundant green fluorescence, which is the mutant huntingtin protein in brain tissue of the mouse model with Huntington's disease. The lower image is that same brain tissue after a CRISPR-based therapy targeting the *HTT* gene encoding huntingtin, showing greatly reduced mutant huntingtin protein (46).

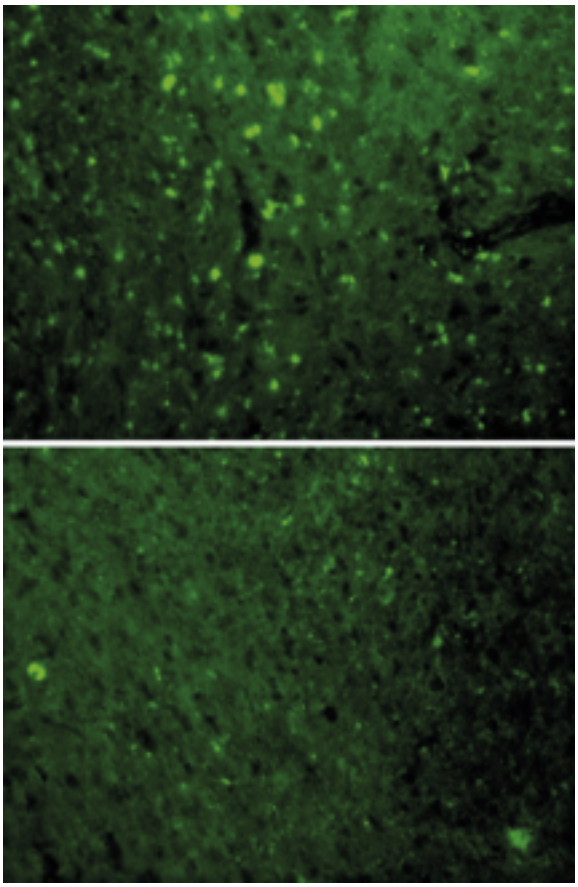

Figure 3. Mutant huntingtin protein decreased in mouse model with CRISPR-based intervention. According to: Eisenstein, Michael (2018); (46).

Muscular Dystrophy-Duchenne Type

Duchenne Muscular Dystrophy (DMD) is an X-linked recessive muscle disease affecting 1 in 5,000 boys with a mutated gene on the X-chromosome inherited from their mothers. It tends to largely be inherited but one-third is from sporadic mutations also. Mutations in the *DMD* gene cause its protein, dystrophin, to lose its shock-absorbing function between muscles and connective tissue around the muscles (47). When dystrophin is working correctly, it attaches to actin and to the connective tissues around the muscle so that the muscle is protected when it

contracts. Mutated dystrophin no longer connects these two entities leading to chronic muscle damage and inflammation just from everyday life and activities (48). Affected children accrue damage to skeletal and myocardial muscle, which at an early age causes them to lose the ability to walk. The course of the disease continues relentlessly, next leading to respiratory and cardiac failure and shortened lifespan. There is no therapy that has succeeded long-term for DMD. Standard treatments have long been corticosteroids and ventilation to help when the diaphragm loses function, and although this has helped to extend DMD patients' lives, it can't save them (48). Lifespan is limited to around 30 years, and is considered the most common fatal Mendelian disorder that is diagnosed in children (49).

The figure (*Figure 4)* shown below illustrates proteins (actin) in the muscle cell attaching to dystrophin, which then is bound to its protein complex in the muscle cell membrane. Normal dystrophin (as shown in *Figure 4*A) bridges protein (actin) to connective tissues around muscle and acts as a shock absorber. However, in the case of DMD seen in *Figure 4*B, dystrophin no longer anchors actin to connective tissues, leading to increased intracellular calcium and dysregulated calcium signaling systems (50).

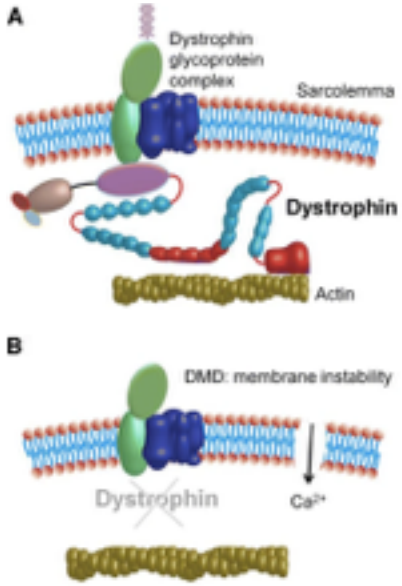

Figure 4. Mutated Dystrophin and loss of muscle membrane stability in Duchenne Muscular Dystrophy. According to: Houang, E.M., et al. (2018); (50).

The dystrophin gene is the biggest gene in the entire human genome, and its protein dystrophin is massive with over 3,000 amino acids. Such a large protein has many locations that mutations can occur and greater than 7,000 mutations have been found in DMD patients (49). Most have a deletion or duplication in at least one exon, although small mutations exist. The dysfunctional dystrophin protein is produced when the number of exons or nucleotides that are deleted cause a reading frameshift, which no longer translates into the functional dystrophin protein (47). Phenotypically, the patient presents with speech delay, decreased muscle function and can be diagnosed in part with the Gower sign, in which the child raising himself from prone position will "walk" his hands back to his legs in order to raise himself up, revealing the weakness of his leg muscles. Serum tests usually show increased muscle creatine kinase (CK) which has spilled into the blood and elevated transaminases. These in combination with muscle weakness make a good case for DMD, although the definitive diagnostic standard is now genetic testing. Genetic tests bypass the invasive muscle biopsies that were done frequently in the past (47).

For DMD, gene transfer therapy is based on removing the mutated gene and replacing it with a gene encoding the normal protein, which then generates normal, healthy dystrophin and stops the degeneration of muscles. The challenge has remained that the DMD gene is massive which would require a huge vector in order to replace it, and our muscles exist all throughout the body meaning a systemic delivery system would be needed. Researchers solved this size and delivery system problem by modifying the large DMD gene into a "micro-dystrophin" gene that could be ferried in the body by using adeno-associated virus (AAV) as a vector for systemic gene transfer. AAV was the ideal vector because unlike adenoviruses, it has a much lower effect on the innate immune response when used in gene therapy (48). AAV has only a small capacity, however, and could not actually carry the entire dystrophin gene. Thus the micro-dystrophin gene was developed to accommodate AAV's size restrictions (51). AAV has become a preferred approach for systemic gene transfer due to its safety and strength in being able to target muscle tissue (51).

The figure (*Figure 5)* below compares normal full-length dystrophin protein (*Figure 5*a) with micro-dystrophin protein produced by the micro-dystrophin gene *(Figure 5*b) (51).

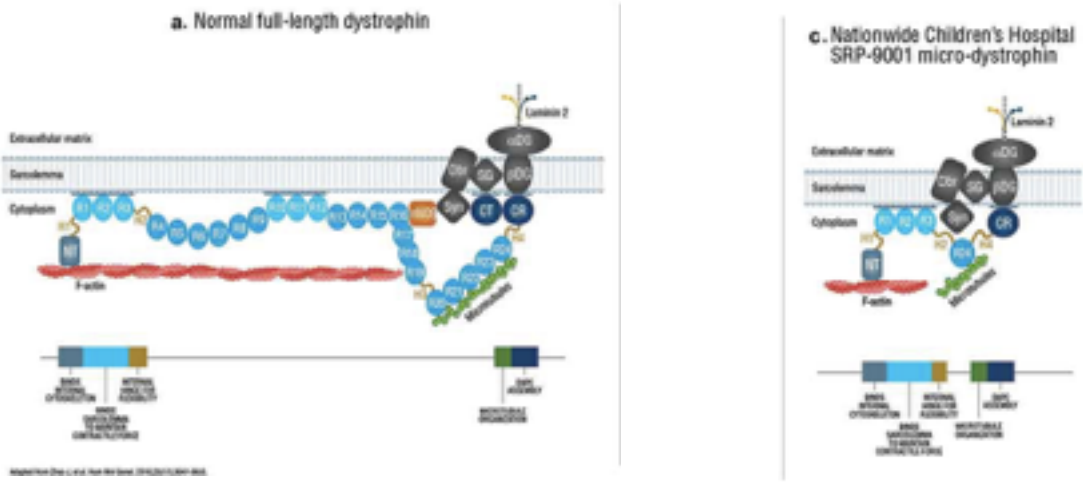

Figure 5. Dystrophin Protein. According to: Asher, Damon R., et al. (51).

The first AAV-micro-dystrophin clinical trial was in 2006 with six boys, ages 5-11, who were injected in the biceps muscle. Results showed only two of the patients had a couple of positive myofibers within 90 days of the injection, and the rest had no response. Further, out of the boys with positive myofibers one had received a low-dose injection and the other a high-dose injection, giving ambiguous results. The biceps injection also meant that only those local fibers could regain any function, not any other muscle fibers in the body. This was not sufficient for DMD patients, whose whole-body musculature was degenerating. In time, much better results have been achieved including the success of the AAV delivery system administered intravascularly in mice. With this mode of administration, there were very positive results that indicated "unequivocal evidence for body-wide improvement" with systemic AAV micro-dystrophin therapy. Clinical trials on humans have since begun with some in late stages now, and the safety of using high-dose AAV vectors is also being tested. High doses are needed for success in generating greater quantities of dystrophin to stop any further muscle damage in DMD patients (51).

The Future of Preventive Genomics

Preventive genomics in its entirety is in its early years, making gains but also going through growing pains to find a place in everyday clinical medicine. From the time the first human genome was sequenced two decades ago, many small steps have unveiled extraordinary findings. There is now an impetus to expand genome sequencing from those high-risk individuals and monogenic diseases to population-wide genomic screening. An Australian study analyzed the impact of population screening on disease prevention and found it could significantly decrease hereditary cancers and mortality, as well as decrease diseases with childhood-onset when compared with current small scale testing (52). With implementation of genomic testing on a greater scale, there lie challenges to decide what specific populations will have access to testing and be actively researched. Thus far, it has been Europeans and people in developed countries who have been sequenced and had their genomes studied overwhelmingly. Moving forward, the concern remains that these populations will amass the benefits in great disproportion to other ethnicities (40).

The full utilization of genomic science will require extensive education for clinicians in order to bring genomic medicine to a clinical forefront. The future demands this, as primary care will have full potential to use genomic information, as long as providers understand and use this in their practices and with their patients. With continued advances to decrease cost and find more rapid sequencing methods, there is conviction that this will increase accessibility and facilitate personalized medicine in all aspects of healthcare (52).

Conclusion

It cannot be understated how the utilization of high-throughput genome sequencing has led to the discovery of vast amounts of genetic information showing gene correlation with disease. The example of genomic sequencing of Mendelian diseases and the plethora of associated variants previously unknown has led to earlier diagnoses and the development of precision therapeutics. Through GWAS, a huge increase in variants have been catalogued and classified, used to connect many previously undiscovered causal genes, their disease associations and phenotypes. Polygenic risk scores are now giving estimations as to the global effect of variants in developing disease, and along with clinical predictors can give a more personalized idea of an individual's overall risk. The dynamic variability of the genome is now firmly cemented by studies showing the re-classification of VUS' which resulted in upgraded and downgraded variant pathogenicity in the time period of a few years. Previous and current research in genome sequencing at all stages of the lifecycle from fetus to child and adult have guided targeted therapeutics to exact genes, allowing "genome editing" and prevention or amelioration of disease. The more diverse, broadened phenotypic data in combination with genetic testing is allowing faster diagnoses in clinical settings and linking of genotypic and phenotypic data has given greater insight into changing phenotypes of disease throughout an individual's lifetime. Genome research is just at the precipice of finding what causes us to get diseases, who gets them and what in our DNA is contributing to it. The antiquated idea of the human genome as static and unchanging has been overturned. Now, knowing the genome is complex and changing with relationships linking genes with other genes, environmental and unknown influences has unlocked a Pandora's box of possibilities. The knowledge gained from laboratories, GWAS', and exhaustive research in the past 20-odd years continues to evolve, change and undoubtedly give people hope that we *can* beat diseases before they begin.

Acknowledgements

I would like to thank my mentor, Prof.dr.sc. Mario Ćuk for inspiring me to take on a topic for my thesis that I knew not about, but which has opened my eyes to the incredible research happening to help people live better lives. Also, thank you Professor for being a role model in the Pediatric field with your energy, positivity and kindness. My parents and grandparents, who saw me through the peaks and valleys of this journey but never wavered from my side and never doubted that I could and would be a doctor someday. And to my rescue pup, Fred, who has been a loyal companion in every season of my life since the day I adopted him in Zagreb, 6 years ago.
It takes a village! Thank you.

All my love,
Annalise

References

1. Feero WG, Guttmacher AE. Genomics, personalized medicine, and pediatrics. Acad Pediatr. 2014;14(1):14-22.
2. Chaudhari, B.P., Manickam, K. & McBride, K.L. A pediatric perspective on genomics and prevention in the twenty-first century. Pediatr Res 87; 2020: 338–344.
3. Green ED, Watson JD, Collins FS. Human Genome Project: Twenty-five years of big biology. Nature. 2015;526(7571):29-31.
4. Levy SE, Myers RM. Advancements in Next-Generation Sequencing. Annu Rev Genomics Hum Genet. 2016;17:95-115.
5. Claussnitzer, M., Cho, J.H., Collins, R. et al. A brief history of human disease genetics. Nature; 2020; 179–189.
6. Scheuerle, Angela E. Which genetic test is best for your patient? AAP News, The Official Newsmagazine for the American Academy of Pediatrics. 2017.
7. Green ED, Watson JD, Collins FS. Human Genome Project: Twenty-five years of big biology. Nature. 2015;526(7571):29-31.
8. National Research Council (US) Committee on Mapping and Sequencing the Human Genome. Mapping and Sequencing the Human Genome. Washington (DC): National Academies Press (US); 1988.
9. Castellanos-Rubio A and Ghosh S. Disease-Associated SNPs in Inflammation-Related lncRNAs.Front. Immunol. 2019; 10:420.
10. Babovic-Vuksanovic D. Genomics as a Scientifically Based Fortune-teller. Mayo Clin Proc.;. 2019; 94(1):7-9.
11. Wise AL, Manolio TA, Mensah GA, et al. Genomic medicine for undiagnosed diseases. Lancet. 2019; 533-540.
12. Clair Morrissey, Rebecca L Walker, The Ethics of General Population Preventive Genomic Sequencing: Rights and Social Justice, The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine. 43(1); 2018: 22–43.

13. Scheuerle, Angela E. Which genetic test is best for your patient? AAP News, The Official Newsmagazine for the American Academy of Pediatrics. 2017.
14. Lalonde E, Rentas S, Lin F, Dulik MC, Skraban CM and Spinner NB. Genomic Diagnosis for Pediatric Disorders: Revolution and Evolution Front. Pediatr.; 2020; 8:373.
15. Delude, C. Deep phenotyping: The details of disease. Nature; 2015; 527 S14–S15.
16. Wise AL, Manolio TA, Mensah GA, et al. Genomic medicine for undiagnosed diseases. Lancet. 2019; 533-540.
17. Mahler EA, Johannsen J, Tsiakas K, et al. Exome Sequencing in Children. Dtsch Arztebl Int. 2019;116(12):197-204.
18. Srivastava S, Love-Nichols JA, Dies KA, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders [published correction appears in Genet Med. 2020 Oct;22(10):1731-1732]. Genet Med. 2019;21(11):2413-2421.
19. Lindstrand et al. From Cytogenetics to cytogenomics : whole-genome sequencing a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. Genome Medicine 2019; 11:68.
20. Lionel, A., Costain, G., Monfared, N. et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. Genet Med 20; 2018; 435–443.
21. Findlay Bewicke-Copley, Emil Arjun Kumar, Giuseppe Palladino, Koorosh Korfi, Jun Wang. Applications and analysis of targeted genomic sequencing in cancer studies. Computational and Structural Biotechnology Journal,Volume 17; 2019; 1348-1359.
22. Gulilat, M., Lamb, T., Teft, W.A. et al. Targeted next generation sequencing as a tool for precision medicine. BMC Med Genomics; 2019; 12, 81.
23. Bamshad, Michael J., Deborah A. Nickerson, Jessica X. Chong. Mendelian Gene Discovery: Fast and Furious with No End in Sight. The American Journal of Human Genetics; 2019; p. 448-455.

24. Wang Bo, Kumar Vivek, Olson Andrew, and Ware, Doreen. Reviving the Transcriptome Studies: An Insight into the Emergence of Single-Molecule Transcriptome Sequencing ; 2019; Front. Genet.
25. Cummings BB, Marshall JL, Tukiainen T, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med.; 2017; 9(386).
26. Johnston J, Lantos JD, Goldenberg A, et al. Sequencing Newborns: A Call for Nuanced Use of Genomic Technologies. Hastings Cent Rep. 2018;48 Suppl 2(Suppl 2):S2-S6.
27. Burstein DS, Gaynor JW, Griffis H, Ritter A, Connor MJO, Rossano JW, Lin KY, Ahrens-Nicklas RC. Genetic variant burden and adverse outcomes in pediatric cardiomyopathy. Pediatr Res. 2021;89(6):1470-1476.
28. Levy B, Burnside RD. Are all chromo-some microarrays the same? What clinicians need to know. Pre-natal Diagnosis. 2019;39:157–164.
29. Loos, R.J.F. 15 years of genome-wide association studies and no signs of slowing down. Nat Commun; 2020; 11, 5900.
30. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2017;45(D1):D896-D901.
31. Rao, S., Yao, Y. & Bauer, D.E. Editing GWAS: experimental approaches to dissect and exploit disease-associated genetic variation. Genome Med; 2021; 13, 41.
32. Daniel J. M. Crouch, Walter F. Bodmer. Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. Proceedings of the National Academy of Sciences; 2020, 117 (32).
33. Wand, H., Lambert, S.A., Tamburro, C. et al. Improving reporting standards for polygenic scores in risk prediction studies. Nature; 2021; 211–219.
34. La Verde, M., De Falco, L., Torella, A. et al. Performance of cell-free DNA sequencing-based non-invasive prenatal testing: experience on 36,456 singleton and multiple pregnancies. BMC Med Genomics 14, 93 (2021).
35. Posey, Jennifer E. Genome sequencing and implications for rare disorders. Orphaned Journal of Rare Diseases; 2019; 14:153.

36. Normand EA, Braxton A, Nassef S, et al. Clinical exome sequencing for fetuses with ultrasound abnormalities and a suspected Mendelian disorder. Genome Med. 2018;10(1):74.
37. Bick D, Jones M, Taylor SL, et al. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseasaes. J Med Genet; 2019;56:783–791
38. Samura, Osamu. Update on noninvasive prenatal testing: A review based on current worldwide research. Obstet. Gynaecol. Res.; 2020; Vol. 46, No. 8: 1246–1254.
39. Pasquale Cacciatore, Laurenske A. Visser, Nasuh Buyukkaramikli, Catharina P. B. van der Ploeg and M. Elske van den Akker-van Marle. The Methodological Quality and Challenges in Conducting Economic Evaluations of Newborn Screening: A Scoping Review. Int. J. Neonatal Screen; 2020; 6(4), 94.
40. Shendure J, Findlay GM, Snyder MW. Genomic Medicine-Progress, Pitfalls, and Promise. Cell. 2019;177(1):45-57.
41. Farnaes, L., Hildreth, A., Sweeney, N.M. et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. npj Genomic Med; 2018; 3, 10.
42. Sabatello M, Appelbaum PS. Honey, I Sequenced the Kids: Preventive Genomics and the Complexities of Adolescence. Am J Bioeth. 2015;15(7):19-21.
43. Lee, J.M. CAG Repeat Not Polyglutamine Length determines timing of Huntington's disease onset. Cell; 2019; pp. 178, 887–900.
44. Halmans, Peter A., Thomas H. Massey, Lesley Jones, Genetic modifiers of Mendelian disease: Huntington's disease and the trinucleotide repeat disorders, Human Molecular Genetics, Volume 26, Issue R2, 01; 2017; Pages R83–R90.
45. Tabrizi, Sarah J., Ghosh, Rita, and Leavitt, Blair R. Huntington lowering strategies for disease modification in Huntington's Dissease. Neuron; Vol 101, Issue 5; 2019; p. 801-819.
46. Eisenstein, Michael. CRISPR takes on Huntington's disease. Nature 557; 2018; S42-S43.
47. Aartsma-Rus A, Ginjaar IB, Bushby K. The importance of genetic diagnosis for Duchenne muscular dystrophy. Journal of Medical Genetics ; 2016;53:145-151.
48. Kupatt, C., Windisch, A., Moretti, A. et al. Genome editing for Duchenne muscular dystrophy: a glimpse of the future?. Gene Ther ;2021.

49. Olson, Eric N. Toward the correction of muscular dystrophy by gene editing. PNAS; 2021; 118 (22).
50. Houang, E.M., Sham, Y.Y., Bates, F.S. et al. Muscle membrane integrity in Duchenne muscular dystrophy: recent advances in copolymer-based muscle membrane stabilizers. Skeletal Muscle; 2018; 8, 31.
51. Damon R. Asher, Khampaseuth Thapa, Sachi D. Dharia, Navid Khan, Rachael A. Potter, Louise R. Rodino-Klapac & Jerry R. Mendell . Clinical development on the frontier: gene therapy for duchenne muscular dystrophy. Expert Opinion on Biological Therapy; 2020 ; pp. 263-274.
52. Babovic-Vuksanovic D. Genomics as a Scientifically Based Fortune-teller. Mayo Clin Proc.;. 2019; 94(1):7-9.

Biography

Annalise Rafaela Bricca is from Portland, Oregon, in the great Pacific Northwest of the United States. She graduated from St. Lawrence University in 2009 with a Bachelor of Arts degree, double majoring in Fine Arts and French. After working in retail and living in France teaching, she took the first step toward her medical career and went back to college at Portland State University for 2 1/2 more years to fulfill all the pre-medical coursework to apply to medical school in the United States. Older and wiser, she realized medical education is excellent in Europe and left the US for Croatia to begin her journey at the University of Zagreb, Medical Studies in English program.

During her 6 years in Zagreb, Annalise was very active in dog rescue, having adopted her dog Fred in Zagreb and going on to foster 15 more dogs, for all of whom she found loving homes. She got her love of people and animals from her kind, generous and warm parents who also have been her greatest supporters through it all.