

## LAB REPORT

# Testing a scalable qPCR chip for early genetic screening of rare diseases

Anika Abbott, and Kaveri Dole ■ Andover High School, Andover, Massachusetts, United States of America

*Reviewed on 9 May 2026; Accepted on 22 June 2026; Published on 26 October 2026*

With over 10,000 known rare diseases, nearly 4% of newborns are estimated to have one, creating a growing diagnostic challenge. Rare disease diagnoses are often delayed until symptoms become apparent, further debilitating patients. Screening for all rare diseases is unfeasible due to financial constraints and limited resources. Current diagnostic strategies rely on symptom-driven genetic testing, which is costly, time-consuming, and impractical for broad screening. To improve early detection, we propose a real-time quantitative Polymerase Chain Reaction (qPCR) chip utilizing TaqMan technology. This approach employs specific primers and fluorescent probes that bind to the DNA and are cleaved during amplification, enabling rapid and accurate screening. Our long-term goal is the ability to screen a single blood sample for hundreds of rare genetic disorders within one scalable qPCR run, provided the causative genes and mutations are known. Although full-scale multiplex screening is our long-term goal, our initial validation experiment outlined in this lab report targets MECP2 to diagnose Rett Syndrome. Limiting initial screening to one gene allowed us to prioritize validating assay performance and optimizing the system before expansion. By converting rare disease diagnostics from symptomatic testing to proactive screening, earlier intervention is enabled.

**Keywords:** Rare disease screening, qPCR, MECP2, Rett syndrome, allele discrimination



*The authors were mentored by Lindsey L'Ecuier. Andover High School, Andover, Massachusetts, United States of America; Peter Horanyi, UCB Pharmaceuticals, Cambridge, Massachusetts, United States of America. Please direct correspondence to: Lindsey L'Ecuier, lindsey.lecuier@andoverma.us. This is an Open Access article, which was copyrighted by the authors and published by BioTreks in 2026. It is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.*

**R**ARE diseases are typically diagnosed through clinical evaluation followed by genetic testing. This type of genetic testing is among the most definitive tools available for diagnosing rare diseases with known genetic causes, yet it is typically only pursued reactively, after symptoms have already intensified. Targeted gene tests or sequencing approaches, including Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS), are time-consuming and costly, causing many patients to remain undiagnosed for prolonged periods of time, a process that carries significant financial burden (Reuter & Kohler, 2019). With clinical WES typically costing \$2,500 and WGS exceeding \$7,000–\$10,000, alongside the hundreds of dollars spent for pre and post test genetic counseling. Despite these excessive costs, oftentimes insurance agencies deny coverage altogether classifying genetic testing as “experimental or investigative” and ignoring the reality that it is often the only path to an

accurate diagnosis (*The Ultimate Guide to Overcoming Genetic Testing Insurance Denials in 2025: Proven Strategies for Patients and Families*, 2025). Additionally, this creates a troubling disparity in access where patients with more financial resources or better insurance plans are far more likely to receive a timely and accurate diagnosis when compared with those without these privileges. The medical consequences of this delay and these disparities are severe and in some cases fatal (*Patients Face Insurance Hurdles for Genetic Testing, UCLA Study Says*, n.d.).

By the time many patients receive an accurate diagnosis, these limitations have already taken a serious toll, including but not limited to organ damage, pulmonary impairment, heart disease and loss of mobility (*The Devastating Impact of a Late Diagnosis in Rare Diseases*, n.d.). This underscores the urgent need for a faster, more affordable, multiplex screening technology.

## Current technologies

Currently, every child born in the United States goes through state-mandated newborn genetic testing for a predetermined list of up to 50 diseases, and about 3,000 newborns test positive for one of these diseases per year (*Newborn Screening - Understanding Genetics - NCBI Bookshelf*, n.d.). This screening is done using Heel Stick Dried Blood Spot (DBS) Testing, which is a minimally invasive method where blood is carefully drawn from the heels of newborns using capillary action so as to not damage the blood sample by squeezing it. Then, drops of this blood are placed and dried onto a DBS card to send to a lab for analysis. The analysis is done using tandem mass spectrometry (MS/MS) and other biochemical assays to detect metabolic and enzymatic abnormalities associated with the target conditions (Wagner et al., 2014). While this approach is effective for the diseases it covers, it is limited to conditions with detectable biochemical markers and can't easily be expanded to the thousands of rare diseases caused by specific genetic mutations (*Newborn Screening - Understanding Genetics - NCBI Bookshelf*, n.d.). As a result, the vast majority of rare genetic diseases fall outside the scope of current newborn screening panels, leaving affected newborns without a diagnosis during critical early windows when intervention would be most beneficial, a gap that the existing newborn screening infrastructure is not currently equipped to fill (*Genetic Testing*, 2026).

Prenatal diagnosis of genetic conditions currently relies on invasive procedures which carry inherent risk to the mother and the fetus. Chorionic Villus Sampling (CVS) involves the insertion of a needle through the abdomen into the placenta to obtain a small blood sample from which DNA can be extracted. Similarly amniocentesis is a technique that requires the insertion of a needle into the amniotic sac to draw amniotic fluid in which the baby's DNA can be found. Both procedures carry risk of miscarriage due to the invasive nature.

A less invasive method exists in the form of cell-free fetal DNA (cffDNA) in which small fragments of fetal DNA continuously shed off the placenta, entering the mother's bloodstream during pregnancy. These fragments can be isolated from a mother's blood draw and analyzed without any direct intervention with potential of harming the fetus. This makes cffDNA an ideal sampling method for prenatal genetic screening. However, while cffDNA provides a noninvasive source of fetal genetic material, identifying disease-causing mutations within these DNA fragments still requires genetic testing. Current cffDNA screening commonly relies on Next-Generation Sequencing (NGS) approaches, including WES and WGS, which remain relatively expensive and unoptimized for large-scale screening of

hundreds of individual rare disease mutations (Zhong & Chiu, 2022).

One promising alternative to reactive genetic testing or screening using NGS from cffDNA samples is the use of real-time quantitative polymerase chain reaction (qPCR), a rapid and cost-effective method for detecting specific genetic variants from a small biological sample. Unlike WES or WGS, which sequence large portions of the genome, qPCR targets known mutations directly, making it faster and significantly cheaper per test. Recent large-scale newborn screening programs have demonstrated that TaqMan-based multiplex qPCR assays can screen for multiple genetic diseases simultaneously from a single dried blood spot sample with high accuracy (Kimizu et al., 2024). This establishes qPCR as a validated foundation for proactive genetic screening. However, existing qPCR-based screening programs remain limited to a small set of well-established target diseases.

## Biological basis for a multiplex screening approach

Although rare diseases are individually uncommon, many share underlying biological pathways meaning that children who have one rare disease often have multiple (Buphamalai et al., 2021). This development is crucial to the screening of rare diseases because it suggests that the testing and potential diagnosis of numerous rare diseases simultaneously is not arbitrary but rather a careful and scientific approach which has a great potential of improving diagnostic efficiency while remaining cost effective.

This connectivity can be seen in the three conditions meticulously selected during our prior study. Congenital Central Hypoventilation Syndrome (CCHS) caused by mutations in *PHOX2B* gene, Autism Spectrum Disorder (ASD) caused by mutations in the *SHANK3* gene, and Rett Syndrome (RTT) caused by mutations found in the *MECP2* gene are all connected through dysfunctions of the Autonomic Nervous System (ANS). The ANS serves as the biological framework guiding future gene selection for the proposed chip. A multigene screening panel could yield clinically meaningful diagnoses to help separate patients into similar biological subgroups.

This current study validated the qPCR assay for the *MECP2* gene, the first step in building this broader multiplex screening system. The same technical framework will later extend to screen for hundreds of rare diseases simultaneously.

## MECP2 & RTT

The target gene for initial experimental validation was *MECP2*. The *MECP2* protein, coded for by the *MECP2*

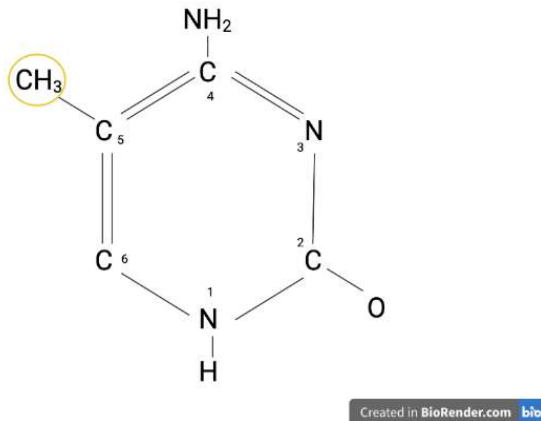


Figure 1. 5-Methylcytosine Structure

gene, is an essential part of embryonic development, the process where a fertilized single-celled zygote goes through rapid cell division, differentiation, and tissue organization to transform the embryo into a human fetus (Khan & Ackerman, 2023). MECP2 specializes in identifying DNA methylation sites, which are specific locations along the DNA strand that have been chemically tagged to regulate gene activity (Du et al., 2015). DNA methylation is an epigenetic mechanism, meaning that it affects how genes are expressed without changing the underlying DNA sequence itself. It involves the transfer of a methyl group with the chemical formula of CH<sub>3</sub> onto the C5 position of a cytosine base, one of the four nucleotide building blocks of DNA. This produces a modified base called 5-methylcytosine, as shown in Figure 1. The addition of the methyl group acts as a molecular switch, either recruiting proteins that shut a gene down or physically blocking transcription factors from accessing that region of the DNA (Moore et al., 2012).

When these transcription factors are blocked or recruited away, the gene they would normally activate is silenced. This directly controls how much mRNA is produced from that gene, and consequently how much of the corresponding protein is produced. In this way, DNA methylation determines which genes are turned off or on in a given cell at a given time, playing a critical role in normal development (Breiling & Lyko, 2015).

MECP2 belongs to a family called methyl-CpG binding domain (MBD) proteins. CpG refers to a region of DNA where a cytosine nucleotide (C) is followed immediately by a guanine nucleotide (G). These sites are common targets for methylation. MBD proteins contain a conserved sequence of approximately 70 amino acids that allow them to selectively recognize and bind to methylated CpG sites. Once bound, MECP2 recruits chromatin remodeling complexes, which are

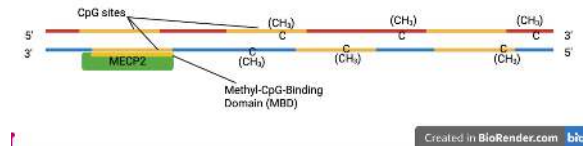


Figure 2. MECP2 protein

large molecular machines that physically reorganize chromatin to compact them further and make local DNA inaccessible to the transcriptional machinery. This silences nearby genes as well.

When the MECP2 gene carries a pathogenic mutation, this entire regulatory process is disrupted. The resulting protein may be unable to bind methylated DNA properly, or may fail to recruit the necessary silencing complexes, causing genes that should be turned off to remain inappropriately active.

These mutations result in the failure to silence genes that should be turned off during neurodevelopment, disrupting the precise timing of brain maturation and leading to the progressive neurological decline characteristics of the disorder (Liyanage & Rastegar, 2018). This demonstrates that precise regulation of MECP2 expression is critical for normal neurodevelopment, and that disruptions in either direction can have significant clinical consequences.

Both loss and overexpression of MECP2 can become disease-causing mutations. For example, MECP2 duplication syndrome, in which an extra copy of the gene causes overexpression of the protein, has also been linked to Autism Spectrum Disorder (ASD) (Nagaran & Hogart, 2006). Conversely, loss-of-function mutations that reduce or eliminate MECP2 activity most commonly produce RTT (Wen & Cheng, 2017). RTT is a congenital condition with varying symptoms including hindered growth and development, loss of bodily functions or control, seizures, scoliosis, and breathing disturbances (*Understanding Rett Syndrome*, n.d.).

Because the genetic basis of RTT is well-established and the MECP2-disease relationship has been extensively characterized by previous studies, it served as an ideal validation target for our qPCR assay. Demonstrating reliable detection of MECP2 mutations provides a proof-of-concept for the broader screening approach before expanding to additional genes with different types of mutations. These include frameshift mutations caused by insertions or deletions rather than single-nucleotide polymorphism (SNP) variations like the RTT-causing mutation in MECP2. Beyond RTT, MECP2 variants have been increasingly documented in individuals with autism spectrum disorder (ASD), making it a gene of high clinical and research relevance

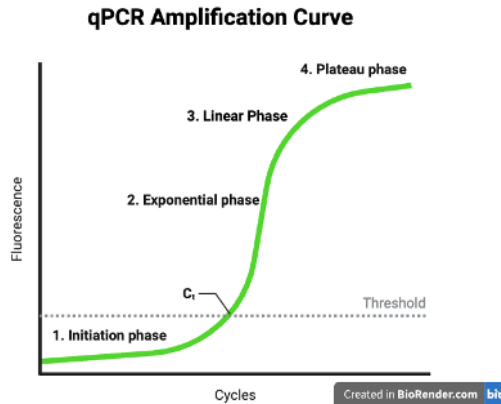


Figure 3. General qPCR Amplification Curve

(Human Gene Module, n.d.).

### qPCR and TaqMan

qPCR is a laboratory technique used to amplify and quantify a DNA sequence in real time. Similar to Polymerase Chain Reaction (PCR), qPCR uses repeated cycles of heating and cooling used to replicate a DNA segment. Unlike standard PCR, qPCR measures the amount of DNA produced after each cycle in real time.

qPCR amplification occurs at four distinct phases that includes an initiation, exponential, linear and plateau phase as seen in Figure 3. The initiation phase does not contain enough DNA for detection to be accurate, since the amount of DNA doubles every cycle, the initial duplication cycles contain too little DNA. This is also called “noise” as it contains irrelevant and unreliable sporadic spurts of detection. During the geometric phase, there is a reagent surplus and sufficient DNA is present for accurate detection, producing the exponential increase in fluorescence. As reagents are used up, the reaction enters the linear phase in which amplification slows down and the curve begins to flatten or plateau. Finally in the plateau phase, reagents are exhausted and all amplification comes to a halt. Since only the exponential phase produces reliable quantitative data, a horizontal fluorescence threshold line is set within that phase in order to determine the Ct value for the reaction.

The Ct value, also referred to as the Quantification cycle (Cq) value, is the cycle number at which the fluorescent signal crosses the threshold and is inversely proportional to the starting amount of DNA. If more DNA is present, fewer cycles are needed to reach or surpass the threshold. Conversely, samples with low DNA concentrations will produce high Ct values. At high Ct values, the signal becomes difficult to distinguish from the background noise, hence why the data above the Ct cutoff is typically considered unreliable.

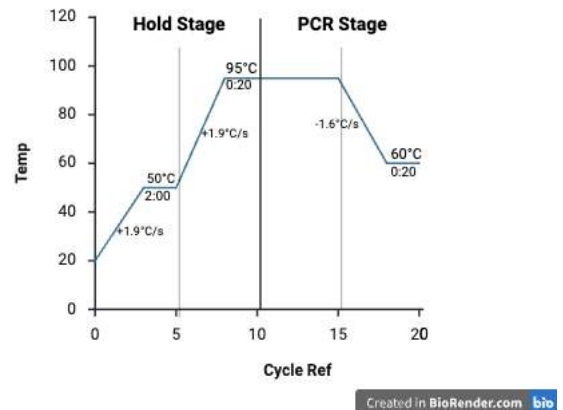


Figure 4. Thermocycling Stages

Ct values are therefore the primary numerical output of a qPCR experiment and form the basis of all analysis.

Thermocycling refers to the process of repeatedly cycling the reaction through a precise series of temperatures, each driving a different step of the PCR process. First, a high temperature of around 95°C causes the double-stranded DNA to denature, meaning the two complementary strands of the DNA molecules are forced apart from each other. This exposes the individual strands so that the primers and probes can access them. The temperature then drops to allow annealing, where the primers and the fluorescent probes bind to their complementary sequences of the now single-stranded DNA template. Finally the temperature is raised slightly to allow the DNA polymerase enzyme to extend the new DNA strand by reading the template and adding matching nucleotides one by one. Each complete cycle of these steps approximately doubles the amount of target DNA present, which is why PCR can generate millions of copies of a specific sequence within just a few hours.

TaqMan based qPCR is a further enhanced technique using short single stranded DNA probes complementary to the original target sequence. The probe has a fluorescent reporter dye on one end and a quencher on the other end which suppresses the fluorescence when the probe is intact. Once the Taq polymerase enzyme has added enough nucleotides to reach the probe, it separates the reporter from the quencher and releases the detectable fluorescent signal. DNA and fluorescence increase proportionally. By designing TaqMan probes which are complementary to the normal or mutant alleles, assays are able to distinguish between the two, confirming the presence or absence of spe-

cific variants. This highly specific method of detection makes TaqMan qPCR ideal for the screening of rare diseases, given the healthy and mutated DNA sequences.

Fluorescein amidite (FAM) and Xanthene dye (VIC) fluorescence were monitored separately across 40 cycles. FAM and VIC are two different fluorescent dyes which we attached to either the mutant or wild-type alleles. Since FAM and VIC emit light at different wavelengths, the qPCR machine can measure each signal independently within the same well, allowing it to determine which allele is present without the two signals interfering with each other. The annealing and extension steps were combined into a single step at a set point of 60°C, which is the temperature at which the probes bind most specifically and selectively to their intended target sequences while minimizing the chance of accidentally binding to the wrong allele.

## Materials and methods

To validate our ability to distinguish wild-type from mutant alleles of the MECP2 gene, our approach used pre-validated, commercially synthesized DNA templates encoding both wild-type and mutant sequences cloned into a standard high-copy bacterial plasmid vector.

These templates served as defined, quantifiable positive controls for the assay, meaning they were samples with a known verified sequence used to confirm that the assay is detecting what it is supposed to detect before testing any unknown samples. Fluorescent probe-based qPCR was then used to distinguish between the two templates using dual-labeled hydrolysis probes specific to the wild-type allele (VIC fluorophore) and the mutant allele (FAM fluorophore), respectively. Fluorophore is a fluorescent chemical compound that emits a specific color of light when activated, allowing each allele to produce its own distinct detectable signal.

Allele discrimination, or the ability of an assay to tell two versions of the same gene apart, was the goal for this validation experiment. It was performed using a pre-validated TaqMan SNP Genotyping Assay (Assay ID C\_356492556\_10; Thermo Fisher Scientific) targeting rsID rs797045692 within the MECP2 gene, the SNP mutation associated with RTT. This dual-probe assay uses a VIC-labeled probe to detect the wild-type allele and a FAM-labeled probe to detect the mutant allele, enabling simultaneous discrimination of both sequences within a single reaction. Since each probe is paired with a different fluorescent dye, the assay can distinguish which version of the gene is present based purely on which color of fluorescent signal is detected.

Wild-type and mutant MECP2 sequences were ob-

tained as synthetic pMA plasmids from Thermo Fisher Scientific. A plasmid is a small, circular piece of DNA that exists separately from a cell's main chromosome and is capable of replicating independently inside of a host cell. pMA is a standard cloning vector from the pMX plasmid family developed by ThermoFisher's GeneArt. pMA carries two important features that make it useful for this purpose. First, it supports high-copy replication in *E. coli*. This means that when the plasmid is introduced to the *E. coli* bacteria, each bacterial cell produces many copies of the plasmid rather than just the one. As the bacteria multiply, so do the plasmids, allowing large quantities of the target DNA sequence to be grown and collected efficiently. Secondly, pMA carries an ampicillin resistance marker, which is a gene that allows bacteria containing the plasmid to survive exposure to ampicillin, an antibiotic that would normally kill them. This acts as a selection tool: when bacteria are grown on ampicillin-containing media, only those that successfully took up the plasmid will survive, ensuring that all surviving bacteria contain the target sequence. Together, these two features make pMA a very reliable and practical way to produce consistent, high-quality DNA templates to use in our qPCR assay.

Essentially, in our experiment, the plasmids act as blood samples, with one plasmid of a healthy patient and one plasmid of a patient with RTT. Since there are legal liabilities associated with obtaining real blood samples, we utilized plasmids to simulate them in our validation. With each of the plasmids, we inserted a target sequence of interest: wild-type MECP2 and the pathogenic variant NM\_001110792.2(MECP2):c.1497A>G (p.Ter499Trp), corresponding to rsID rs797045692. This notation c.1497A>G describes the exact location and nature of the mutation at the DNA molecular level, where an adenine (A) nucleotide at position 1497 of the coding sequence is replaced by a guanine (G). The notation p.Ter499Trp describes the consequence of this mutation at the protein level, where the normal stop codon at amino acid position 499 that signals the cell to stop building the protein, is replaced by a tryptophan codon, causing the protein to continue being assembled incorrectly beyond its intended endpoint. This produces an abnormal extended protein that disrupts normal MECP2 function.

Plasmids were resuspended in TE buffer which is a standard solution of Tris and EDTA used to stabilize and preserve DNA. Then the plasmids were diluted to a working concentration of approximately 1–2 ng/μL (nanograms of DNA per microliter of solution) in nuclease-free water prior to use.

## qPCR reaction setup

Reactions were prepared using TaqMan Fast Advanced Master Mix (Catalog No. 4444556; Thermo Fisher Scientific) following the manufacturer's recommended protocol for standard 96-well plates, which specifies a total reaction volume of 18  $\mu$ L per well. The master mix included most of the chemical components needed for a PCR reaction, including the DNA polymerase enzyme that copies the DNA, the nucleotides that serve as building blocks for new DNA strands, and the buffer chemicals that maintain stable reaction conditions.

Rather than preparing each reaction individually, a bulk stock solution was prepared at sufficient volume to fill seven wells, with a small excess to account for pipetting loss. This stock solution consisted of 70  $\mu$ L of 2X (twice the working concentration which becomes the correct strength once diluted by the other reaction components) master mix, 7  $\mu$ L of 20X TaqMan assay, and 49  $\mu$ L of nuclease-free water. This water is highly purified from which all enzymes that could degrade DNA have been removed, used here to bring the total reaction volume to the correct amount without introducing any contaminating material. The TaqMan assay was originally supplied as a 40X concentrated stock and was first diluted to a 20X by mixing an aliquot with an equal volume of TE buffer before being added to the stock solution.

17  $\mu$ L of this stock solution was then distributed equally into each of the six wells. Finally, 1  $\mu$ L of template DNA was added individually to each well, with wild-type MECP2 DNA added to three wells and mutant MECP2 DNA to the other three. This brought the total volume in each well to 18  $\mu$ L. The result was three replicate reactions for each template, which is standard practice to allow for the calculation of reliable averages and the identification of any inconsistent results caused by technical error.

It should be noted that the template DNA was added as a last step directly to individual wells rather than being incorporated into the bulk stock solution. This likely introduced variability between replicates, as manually pipetting 1  $\mu$ L into individual wells is particularly susceptible to small pipetting errors at such low volumes. This is the most probable explanation for the variation observed in our Ct values across replications, and future experiments would benefit from incorporating the template DNA into the bulk stock solution before distribution to improve consistency.

All reactions were run on a Bio-Rad CFX Connect Real-Time PCR System (Bio-Rad Laboratories) using a standard 96-well plate (HSP-9601; Bio-Rad), and Ct values were recorded for each well. The plate includes small individual wells arranged in an 8-by-12 grid, each capable of holding a small volume of liquid

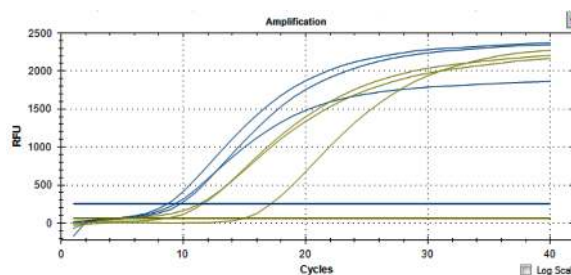


Figure 5. Results Graph

independently, allowing many different reactions to be run simultaneously under identical conditions. The Bio-Rad CFX Connect is a qPCR machine, meaning it not only cycles through the temperatures required for PCR but also measures fluorescent signals from each well after every cycle, allowing the accumulation of DNA to be monitored as it happens rather than only at the end. Thermocycling conditions followed the manufacturer's recommended protocol for TaqMan Fast Advanced Master Mix.

## Data analysis

Raw Ct values were exported and processed in Microsoft Excel. Well A03, which yielded a Ct value of 14.86 — substantially higher than the other VIC replicates (5.23 and 8.34) — was identified as an outlier and excluded from analysis. With a mean Ct value of 6.79 and a standard deviation of 1.56, the A03 value falls more than 2 standard deviations above the mean Ct value classifying it as an outlier. Mean Ct values and standard deviations were calculated for the remaining replicates of each condition. Results are presented as a bar graph with custom error bars representing the standard deviation of replicate measurements.

## Results

All six wells produced amplification curves consistent with successful PCR as shown in Figure 5, showing fluorescence rising through an exponential phase before leveling off into a plateau. No cross-channel signal contamination was visually observed, suggesting the FAM and VIC probes were functioning independently within each well.

Raw Ct values for the mutant replicates (FAM channel, wells B01-B03) were 9.80, 8.75, and 9.39, yielding a mean Ct of 9.31. Raw Ct values for the wild-type replicates (VIC channels, wells A01-A03) were 5.23, 8.34, and 14.85. Well A03, which returned a Ct of 14.86, was identified as an outlier. It fell far above the other two WT replicates and is most likely the result of a pipetting error during the 1  $\mu$ L template addition step. Manu-

ally pipetting such a small volume into individual wells introduces variability, as even a slight inaccuracy can meaningfully alter the amount of DNA in a reaction. Well A03 was therefore excluded from further analysis. The two retained WT replicates yielded a mean Ct of 6.78.

While both templates amplified successfully, the results of this experiment had two important limitations that motivated a second run. First, excluding the outlier left only two usable WT replicates instead of three, which is insufficient for calculating a statistically reliable mean and standard deviation (Serdar et al., n.d.). Second, the experiment was designed with each well containing only one probe (either FAM or VIC), rather than both probes simultaneously in the same well. True allele discrimination which is the purpose of our experimentation, requires that both probes compete within a single reaction well so the instrument can determine which allele is present based on which fluorescent signal dominates. Running the probes separately does not replicate this condition and therefore does not fully validate the assay's ability to distinguish alleles in a real diagnostic scenario.

To address these limitations, a second, more rigorous experiment was performed using a QuantStudio 12K Flex instrument (Applied Biosystems) with a 384-well plate format. This instrument is capable of monitoring multiple fluorescence channels simultaneously within the same well, making it better suited for the dual-probe allele discrimination the assay requires. The same TaqMan SNP Genotyping Assay and synthetic plasmid templates were used as in the previous experiment. Each template condition (both MUT and WT) was run in five replicate wells, each containing both the FAM-labeled mutant probe (Target 1) and the VIC-labeled wild-type probe (Target 2) simultaneously, as is standard for a genotyping assay (Buphamalai et al., 2021). This meant that in each well, both probes were present and competing, so the instrument could determine allele identity based on which signal appeared earlier and stronger.

Results were highly consistent across replicates. For the mutant template, the FAM (mutant-specific) probe produced a mean Ct of 9.923, reflecting early, robust amplification as seen in Figure 6. In the same wells, the VIC (wild-type-specific) probe produced a mean Ct of 10.512, a later weaker signal indicating low-level non-specific probe activity rather than true allele detection.

For the wild-type template, the VIC probe produced a mean Ct of 8.679, indicating early and strong amplification of the correct allele. In those same wells, the FAM probe produced a mean Ct of 17.849, a very late signal indicating that the mutant-specific probe barely amplified when presented with wild-type DNA as in

Figure 7.

The difference in Ct between the correct probe and the incorrect probe is the key indicator of allele discrimination. A large  $\Delta$ Ct between the intended and unintended probe signals in the same well means the assay is correctly favoring the matching allele. For the WT template, the  $\Delta$ Ct between FAM (wrong probe) and VIC (correct probe) was approximately 9.2 cycles which was a large and meaningful difference indicating that the VIC probe strongly outperformed the FAM probe on wild-type DNA, as expected.

In both template conditions, the probe matching the template allele produced a lower Ct than the mismatched probe in every replicate well, confirming that the assay correctly identified which allele was present in all ten reactions. The small standard deviations across five replicates per condition, where the Ct standard deviation was found to be less than 0.486, further demonstrates that the assay performed with high reproducibility.

Taken together, the results of Experiment 2 demonstrate that the TaqMan SNP Genotyping Assay successfully discriminates between wild-type and mutant MECP2 sequences with high reproducibility (the maximum Ct standard deviation is 0.486 across all conditions) and consistent allele-specific signal directionality across five replicates. These results validate the assay as a reliable tool for MECP2 genotyping and establish a strong proof of concept for incorporating this target into a broader multiplex rare disease screening platform.

## Safety

All procedures were conducted in compliance with institutional biosafety guidelines. Experiments were performed in a standard molecular biology laboratory at the University of Massachusetts Lowell (UML) Center for RNA Research Facilities. No live organisms or hazardous chemicals were used. Synthetic plasmid DNA and PCR reagents were handled according to manufacturer safety data sheets and disposed of in accordance with institutional waste protocols.

Nitrile gloves, laboratory coats, and eye protection were worn throughout all experimental work. Additionally, reagents were stored according to the manufacturer specifications and all waste was separated and disposed of in a biohazard container.

## Next Steps: Validating More Genes

Having successfully validated our qPCR assay with MECP2, the next step is to validate our other two target genes, PHOX2B and SHANK3. Each will undergo a similar validation process, using custom-designed primers

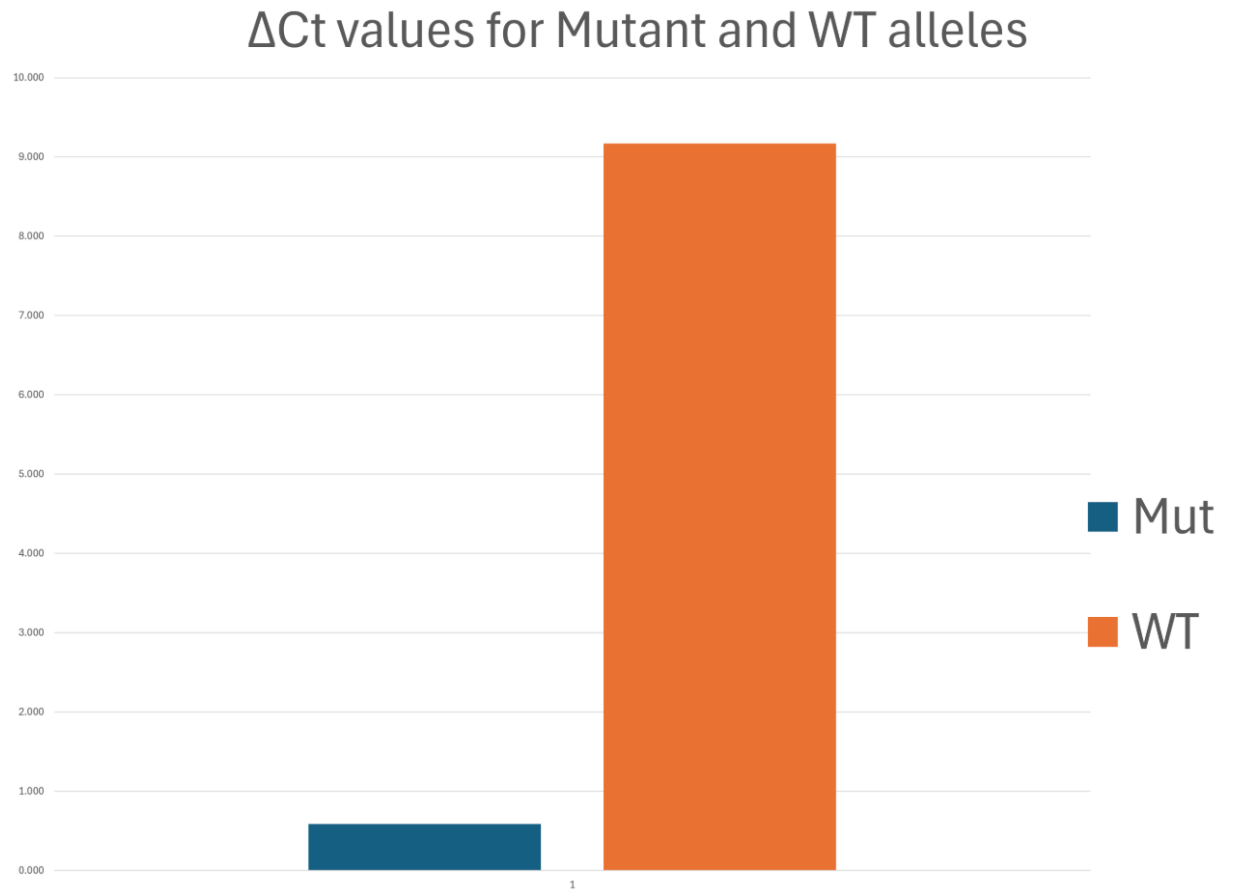


*Figure 6. Wild Type Template Results*

and TaqMan probes specific to the wild-type and mutant sequences of each gene. Assay performance will be evaluated based on amplification efficiency and the separation in Ct values between wild-type and mutant sequences. A clear differentiation between the two confirms that the assay can reliably distinguish a disease-causing mutation from a normal sequence which is the fundamental requirement before any gene can be incorporated into a broader screening panel. Once both genes are successfully validated, this work will directly support our expansion toward a fully multiplexed rare disease screening tool.

Not all disease-causing mutations are equally straightforward to detect, and this becomes an important practical consideration as we scale up. The MECP2 assay was able to use a premade probe because it targets a SNP at a known position. Frameshift mutations, like those found in PHOX2B, are structurally different and require fully custom-designed primers and probes to detect accurately. This added complexity makes the validation process considerably more

expensive than it was for MECP2. We are currently awaiting additional funding through grants, which will allow us to cover these development costs and move forward with our second round of validations.

The ultimate goal of this project is to develop a chip capable of screening for multiple rare genetic diseases simultaneously, from a single patient sample in a single qPCR run. To achieve this without signal interference between targets, each gene must be paired with its own distinct fluorescent reporter dye — such as FAM, VIC, or other available channels — so that every target generates a uniquely identifiable signal within the same reaction. This multiplexing approach is what makes simultaneous multi-disease screening practical, allowing us to maximize the diagnostic information obtained from one sample while keeping the process efficient and scalable. Since qPCR instruments have a limited number of fluorescence channels, targets will also be separated across distinct wells on the chip. Meaning, both will be assigned a fluorescent dye and a physical position of the chip to enable multiplex scaling as

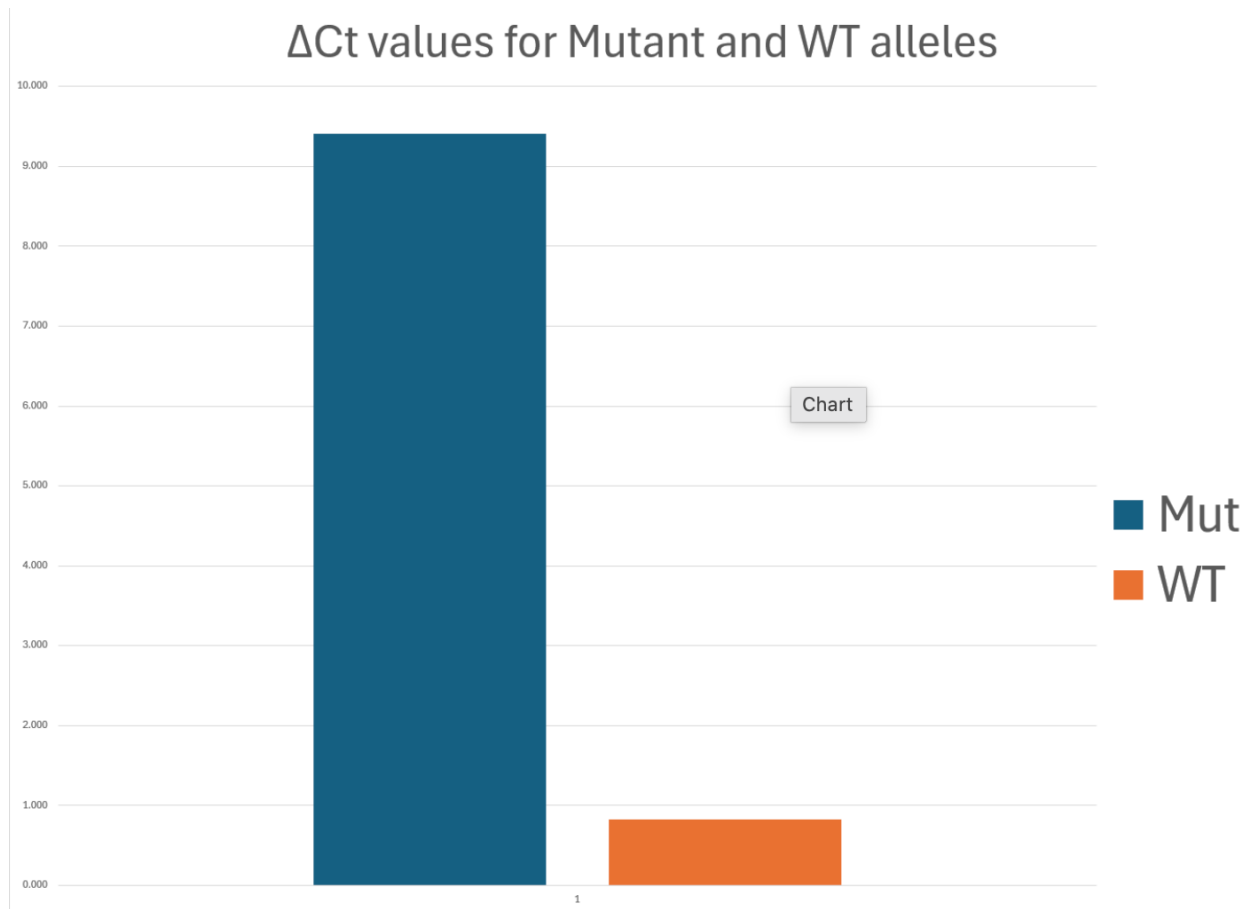


Figure 7. Mutant Template Results

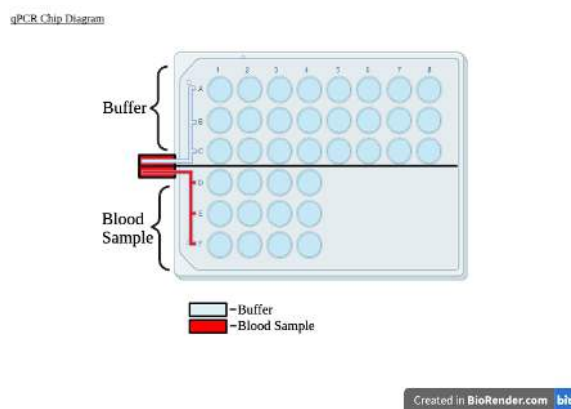


Figure 8. qPCR Chip Design

shown in Figure 8.

### Current technologies

High-throughput qPCR technologies, such as ThermoFisher's OpenArray™ system, demonstrate the

feasibility and power of multiplex genetic testing. OpenArray™ contains a microscope-slide-sized plate containing 3,072 reaction wells, enabling thousands of simultaneous qPCR reactions on a single plate. This platform enables flexible gene expression, microRNA (miRNA), and genotyping assays, and is capable of efficiently handling large sample volumes. However, despite its technical strengths, OpenArray™ is designed for research and is labeled “For Research Use Only - Not for diagnostic procedures” (OpenArray® Technology Overview, n.d.). The system requires trained laboratory personnel, a controlled research setting, and assay plates purchased in bulk, increasing cost and limiting accessibility.

These constraints mean that OpenArray™ and similar platforms are not practical for routine clinical screening or newborn testing, particularly for rare diseases. In contrast, our proposed chip is designed specifically for clinical newborn screening targeting a panel of rare diseases at a smaller scale, lowering the cost per sample while requiring minimal specialized infrastructure and meeting diagnostic grade stan-

dards.

## Author contributions

A.A. and K.D. equally contributed to the research, experimentation, and writing of the manuscript.

## Acknowledgements

We are deeply grateful to Dr. Lindsey L'Ecuyer, whose guidance made this study possible. Dr. L'Ecuyer supported us throughout every stage of this project — from initial research and experimental design through writing, editing, and revision — dedicating countless hours to sharing her knowledge and providing essential feedback that significantly strengthened this manuscript. She also introduced us to the BioBuilders community by establishing our school's BioBuilders club and encouraging us to submit to BioTreks. Through the BioBuilders Final Assembly, we were connected with our mentor Dr. Peter Horanyi of UCB Pharmaceuticals in Cambridge, Massachusetts, whose extensive expertise in qPCR was invaluable in guiding the technical setup of our experiments. We would also like to thank Jack Lepine for his assistance conducting experiments at the University of Massachusetts Lowell Core Research Facilities.

## References

- Breiling, A., & Lyko, F. (2015, July 21). *Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond*. Springer Nature Link. Retrieved May 9, 2026, from <https://link.springer.com/article/10.1186/s13072-015-0016-6>
- Buphamalai, P., Kokotovic, T., Nagy, V., & Menche, J. (2021, November 9). *Network analysis reveals rare disease signatures across multiple levels of biological organization*. National Library of Medicine. Retrieved May 29, 2026, from <https://pubmed.ncbi.nlm.nih.gov/34753928/>
- Du, Q., Luu, P.-L., Stirzaker, C., & Clark, S. J. (2015, April 30). *Methyl-CpG-binding domain proteins: readers of the epigenome*. National Library of Medicine. Retrieved May 9, 2026, from <https://pubmed.ncbi.nlm.nih.gov/25927341/>
- Genetic testing*. (2026, February 27). Mayo Clinic. Retrieved May 7, 2026, from <https://www.mayoclinic.org/tests-procedures/genetic-testing/about/pac-20384827>
- Human Gene Module / Chromosome X / MECP2*. (n.d.). SFARI GENE. Retrieved May 9, 2026, from <https://gene.sfari.org/database/human-gene/MECP2>
- Khan, Y. S., & Ackerman, K. M. (2023, April 17). *Embryology, Week 1 - StatPearls - NCBI Bookshelf*. NCBI. Retrieved May 9, 2026, from <https://www.ncbi.nlm.nih.gov/books/NBK554562/>
- Kimizu, T., Nozaki, M., & Okada, Y. (2024, February 28). *Multiplex Real-Time PCR-Based Newborn Screening for Severe Primary Immunodeficiency and Spinal Muscular Atrophy in Osaka, Japan: Our Results after 3 Years*. National Library of Medicine. Retrieved May 9, 2026, from <https://pmc.ncbi.nlm.nih.gov/articles/PMC10970021/>
- Liyanage, V. R. B., & Rastegar, M. (2018, February 5). *Rett Syndrome and MeCP2 - PMC*. PMC. Retrieved May 9, 2026, from <https://pmc.ncbi.nlm.nih.gov/articles/PMC5798978/>
- Moore, L. D., Le, T., & Fan, G. (2012, July 11). *DNA Methylation and Its Basic Function*. Neuropharmacology. Retrieved May 9, 2026, from <https://www.nature.com/articles/npp2012112>
- Nagarajan, R., & Hogart, A. (2006, June 19). *Reduced MeCP2 Expression is Frequent in Autism Frontal Cortex and Correlates with Aberrant MECP2 Promoter Methylation*. Taylor & Francis Online. Retrieved May 9, 2026, from <https://www.tandfonline.com/doi/abs/10.4161/epi.1.4.3514>
- Newborn Screening - Understanding Genetics - NCBI Bookshelf*. (n.d.). NCBI. Retrieved May 9, 2026, from <https://www.ncbi.nlm.nih.gov/books/NBK132148/>
- Patients face insurance hurdles for genetic testing, UCLA study says*. (n.d.). UCLA Health. <https://www.uclahealth.org/news/article/patients-insurance-genetic-testing-study>
- Reuter, C. M., & Kohler, J. N. (2019, September 3). *Yield of Whole Exome Sequencing in Undiagnosed Patients Facing Insurance Coverage Barriers to Genetic Testing*. National Library of Medicine. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6901723/>
- Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (n.d.). *Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7745163/>
- The Devastating Impact of a Late Diagnosis in Rare Diseases*. (n.d.). Mexico Business. <https://mexicobusiness.news/health/news/devastating-impact-late-diagnosis-rare-diseases>
- The Ultimate Guide to Overcoming Genetic Testing Insurance Denials in 2025: Proven Strategies for Patients and Families*. (2025, May 9). Counterforce Health. Retrieved May 29, 2026, from <https://www.counterforcehealth.org/post/the-ultimate-guide-to-overcoming-genetic-testing-insurance-denials-in-2025-proven->

- strategies-for-patients-and-families/  
*Time to diagnose a rare disease: managing medical uncertainty. A qualitative study.* (2024, August 14). PMC. Retrieved May 7, 2026, from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11323401/>
- Understanding Rett Syndrome.* (n.d.). International Rett Syndrome Foundation. Retrieved May 9, 2026, from <https://www.rett syndrome.org/about-rett-syndrome/understanding-rett-syndrome/>
- Wagner, M., Tonoli, D., Varesio, E., & Hopfgarten, G. (2014, September 22). *The use of mass spectrometry to analyze dried blood spots.* National Library of Medicine. Retrieved May 9, 2026, from <https://pubmed.ncbi.nlm.nih.gov/25252132/>
- Wen, Z., & Cheng, T.-L. (2017, August 3). *Identification of autism-related MECP2 mutations by whole-exome sequencing and functional validation.* Springer Nature Link. Retrieved May 9, 2026, from <https://link.springer.com/article/10.1186/s13229-017-0157-5>
- Zhong, L. P. W., & Chiu, R. W.K. (2022, March 22). *The Next Frontier in Noninvasive Prenatal Diagnostics: Cell-Free Fetal DNA Analysis for Monogenic Disease Assessment.* Annual Reviews. Retrieved May 29, 2026, from <https://www.annualreviews.org/content/journals/10.1146/annurev-genom-110821-113411?>