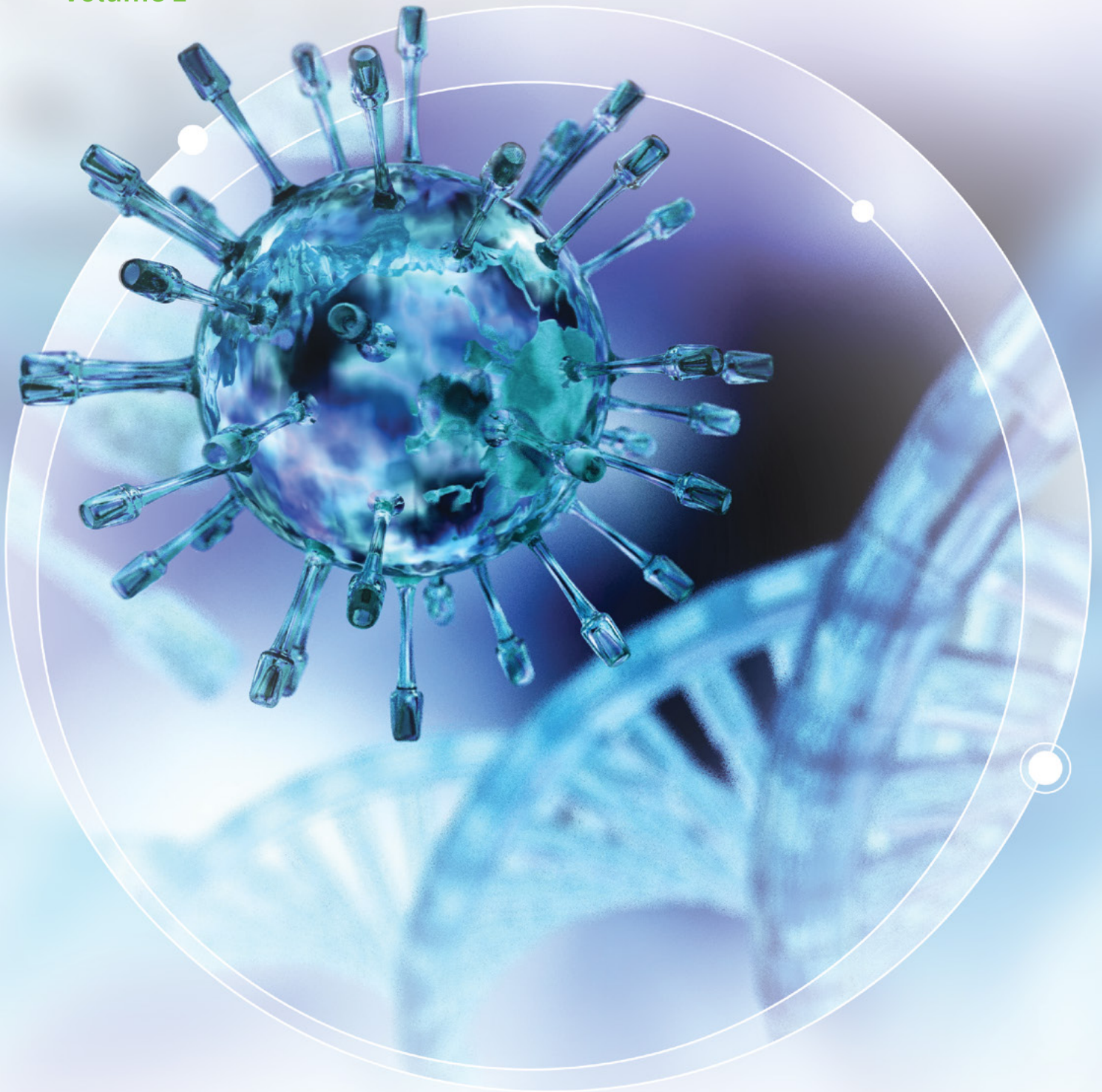


# Gene Therapy and Oligonucleotide Compendium

Volume 1



# Gene Therapy and Oligonucleotide Compendium

## Contents

Introduction	4
Growing trends and emerging technologies for oligonucleotide and gene therapies	6
The global gene therapy & oligonucleotide market	8
Regional landscape	12
North America	12
EMEA (Europe, Middle East, Africa and India)	13
China	14
Japan	15
Korea and South East Asia	16
Modalities	18
Viral Gene Therapy	18
Oligonucleotides	18
CRISPR	19
Technical note summaries	20
Capsid protein analysis	28
Technical Note 1: Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins Using CE-SDS Method	29
Technical Note 2: Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins Using CE-LIF Technology	36
Technical Note 3: Sensitive AAV Capsid Protein Impurity Analysis by CE Using Easy to Label Fluorescent Chromeo Dye P503	42
Nucleic acid sizing analysis	46
Technical Note 4: Accurate Analysis of Double Stranded DNA over an Extended Size Range	47
Technical Note 5: Method Evaluation for RNA Purity Analysis Using CE-LIF Technology	58
Technical Note 6: RNA Analysis for CRISPR by Capillary Electrophoresis with Laser-Induced Fluorescence Detection	66
Technical Note 7: Highly Precise DNA Sizing on the SCIEX Genetic Analysis System	75



<b>Plasmid analysis</b>	<b>79</b>
Technical Note 8: Simple, Fast and Robust Method for Plasmid Purity Testing and Degradation Monitoring	<b>80</b>
Technical Note 9: A Sensitive and Robust Plasmid Analysis Method by Capillary Electrophoresis-Laser Induced Fluorescence	<b>89</b>
<b>Selection of aptamers</b>	<b>93</b>
Technical Note 10: CE-SELEX: Rapid Aptamer Selection Using Capillary Electrophoresis	<b>94</b>
<b>Nucleic acid sequencing</b>	<b>104</b>
Technical Note 11: Bacterial Artificial Chromosome (BAC) End Sequencing Analysis	<b>105</b>
Technical Note 12: Improved Sequencing of Plasmids on the SCIEX Genetic Analysis System	<b>111</b>
Technical Note 13: Method for Long Sequencing Read Lengths on the SCIEX Genetic Analysis System	<b>116</b>
<b>SNP analysis &amp; genotyping</b>	<b>122</b>
Technical Note 14: SNPs, Mutations, and DNA Sequence Variation Analysis Using the GeXP SNPStart Kit	<b>123</b>
Technical Note 15: Multiplex SNP Analysis: Screening Factor V R506Q (Leiden) Mutations	<b>132</b>
Technical Note 16: Multiplex Mutation Genotyping for Human Diseases: Breast Cancer and Familial Mediterranean Fever	<b>141</b>
<b>Gene expression profiling</b>	<b>148</b>
Technical Note 17: Quantitative Multiplexed Gene Expression: An Analysis of Sensitivity, Accuracy and Precision	<b>149</b>
Technical Note 18: Monitoring Multiple Pluripotency Biomarkers after Delivery of Reprogramming Factors to Human Somatic Cells	<b>156</b>
Technical Note 19: Confirming Gene Mutation by CRISPR-Cas9 at the Protein Level and Identifying Proteome-Wide Changes	<b>169</b>
<b>Analysis of oligonucleotide therapeutics</b>	<b>174</b>
Technical Note 20: High Resolution Analysis of Synthetic Oligonucleotides	<b>175</b>
Technical Note 21: Extending the Lower Limits of Quantification of a Therapeutic Oligonucleotide Through Microflow LC-MS/MS	<b>181</b>
Technical Note 22: Quantification of Large Oligonucleotides using High Resolution MS/MS on the TripleTOF® 5600 System	<b>188</b>

# Introduction

**The emergence of new therapeutic modalities facilitates a dramatic increase in the ability to treat a wide range of previously “untreatable” diseases. This increase does not occur in a linear fashion. As scientists overcome the technological hurdles of developing new biologics into safe and effective therapies, more and more diseases become treatable. (Figure 1). This phenomenon continues to hold true for the emergence of oligonucleotide and gene therapies.**

There are over 7000 genetic diseases that could potentially be cured using gene therapy. Rare metabolic diseases, autoimmune disorders, cardiovascular disease and cancers are some of the top disease classes that are addressable with gene therapies. These are diseases that are caused by mutations in one or more genes. There are currently over 1000 clinical trials, in various stages, that involve gene therapies or oligonucleotides. The potential benefit to human health is profound.

This new generation of therapeutic modalities presents new and unprecedented technological challenges in bringing therapies to patients. These challenges have impacted both the speed and ultimate cost of bringing new therapies to market with many gene therapeutics now categorized as some of the most expensive medications in existence today.

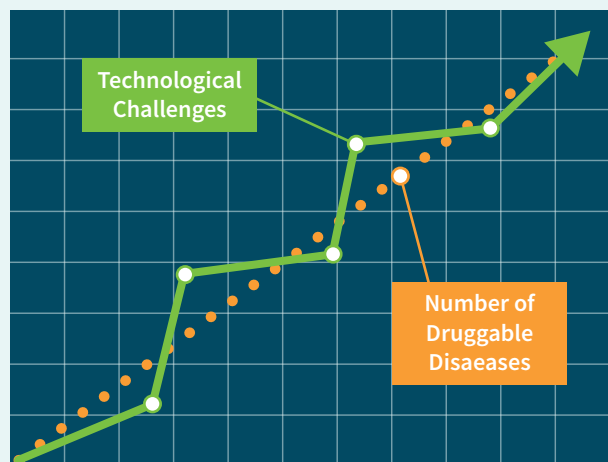


Figure 1

The high cost of gene therapies to the developer and the patient is in part due to development and manufacturing processes that are not well optimized. Much of the analytical testing and characterization of the therapeutic is done at the end of the development process with no means for course correction along the way. Thus, yields are very low – not necessarily the yield within a batch, but rather the number of batches it takes to yield a successful product. Ultimately, the lack of sufficient tools has not only impacted gene therapy programs but also the market as a whole as scientists search for faster and more accurate methods for characterizing these new classes of drugs.



This compendium aims to assist scientists in the characterization of oligonucleotide-based and protein-based starting materials, side reactions and the final products resulting from the design and manufacture of gene and cell therapies. It will also help in quality control efforts by presenting workflows and approaches

to purity testing, sizing, sequencing and expression analysis. Additionally, this compendium contains strategies for gaining a better understanding of the biological impact of genetic mutations and specific gene editing events.





# Mani Krishnan

Vice President and General Manager,  
CE and Biopharma Business Unit, SCIEX

## Growing trends and emerging technologies for oligonucleotide and gene therapies

**While the promise and investment in the development of gene therapies has been high, the technologies for manufacturing and analysis of gene therapy products are still evolving.**



The current state of gene therapy development and manufacturing are similar to the early days of protein therapeutics development and production. There is clearly an expectation that we will follow a model in maturing gene therapy production analogous to how it was done with protein therapeutics, but at a much faster pace. There is a roadmap and, while different, it's not completely unexplored territory. Nonetheless, there are some fundamental differences between production of protein therapies and gene therapies. While it is true that there will be minor post-translational modifications with protein therapeutics, a well characterized master cell bank is enormously beneficial in ensuring that a reliable and well characterized product can be manufactured. With viral vectors, there is, currently, no stable cell line that produces the same vector. The need for each cell to be transfected by plasmid complexes in order to produce the viral vector increases the probability of inducing product heterogeneity. Robust and scalable downstream processes are needed to be able to reproducibly provide a homogenous pool of viral vectors that meet the safety, strength, identity and purity requirements. While there are efforts

*“We are just at the very early stages of the gene therapy industry, where we have a biologics model for development but with many unique challenges.”*

underway to move to stable and immortalized, producer cells, most products in the clinic or that are in the market today still use transient transfection for producing viral vectors.

One of the main challenges in the development of robust and scalable upstream and downstream process for gene therapy products is the turn-around time it takes for the analytical results to be obtained. For example, in gene therapy process development, knowing early on that there is good transfection efficiency is crucial to making sure the viral vector has the right transgene in it that will eventually have the desired efficacy. Today, these assays are cell-based which can take up to 3 weeks to return results. Likewise, ensuring the integrity of capsid proteins for adeno-associated virus (AAV) products is most often accomplished through infectivity studies which can take up to a week. Additionally, high variability, sometimes as high as +/-50% is another unfortunate aspect of many current gene therapy analytical methods.

Assays with faster turn-around time and reliable results lead to better informed decisions to be made quickly about how the process is developed and scaled-up. Similarly, during manufacturing, whether for clinical trials or for approved product, near-real-time, reliable assays would allow faster decisions on whether or not to forward-progress a batch. If any problems were encountered, the process or batch could be aborted earlier saving a great deal of time and money. We are just at the very early stages of the gene therapy industry, where we have a biologics model for development but with many unique challenges. Having robust assays that are tailored to the analysis of gene therapy products, that can provide faster results, and with better accuracy and reproducibility than what is currently available, will allow true acceleration and progression in the market.





# Baburaj Kunnummal

Senior Global Market Manager  
CE & Biopharma Business Unit

## The global gene therapy & oligonucleotide market

**The global growth in the gene therapy market is driven by the increase in the number of clinical trials, government and private funding and an increasing number of partnerships between small biotech and large pharmaceutical companies. Across the globe, 11 gene therapy drugs are on the market.**



In the US alone, the FDA estimates that companies will file more than 200 new investigational drug applications (IND) for gene therapy products by the year 2020. When these are added to the number of applications currently on file, the expectation is for 10-25 new gene therapy drug approvals by the year 2025.

The global gene therapy market size is estimated to be more than \$1 billion as of 2019 and is expected to grow at a compound annual growth rate (CAGR) of 32%. An increase in research and development funding and innovations in gene therapy related research are some of the factors driving the growth in this therapeutic area.

For oligonucleotides, the therapeutic market can be segmented into antisense oligonucleotide (ASO), aptamers, micro RNA (miRNA) small interfering RNA (siRNA). ASO and RNAi accounts for 57% of the oligonucleotide therapeutic development market. There are more than 173 ongoing oligonucleotide therapeutics clinical trials, with

*“The United States holds more than 50% of the oligonucleotide therapeutics market share. Germany and China lead their respective regions in the therapeutic oligonucleotide market.”*





23 molecules in Phase III or later, sponsored by more than 35 companies. There are eight approved oligonucleotide therapeutics on the market. A recent market analysis predicts that the global ASO and RNAi therapeutics market will reach \$4.6 billion by 2022. The United States holds more than 50% of the oligonucleotide therapeutics market share. Germany and China lead their respective regions in the therapeutic oligonucleotide market. The overall growth in the market segment is driven by pharmaceutical company investments in exploring the therapeutic applications of oligonucleotides. Other drivers are the additional targets characterized from the sequencing of the human genome, new generations of antisense oligonucleotides and the development of double-stranded siRNA.

Gene therapy development, including oligonucleotides offers tremendous opportunities for analytical instrument makers and industry leaders to collaborate to design novel analytical tools that can help faster development of these lifesaving therapies.

**References:**

1. Growth Insights on the Global CRISPR/Cas9 Tools Market, 2017 (Frost & Sullivan)
2. Medicines in Development, 2018 Report
3. Global Cell Therapy Market (2017 – 2021), Technavio.com
4. Gene Therapy 2017: Progress and Future Directions, Clin. Transl. Sci. (2017) 10, 242–248
5. Gene therapy on the move, EMBO Mol Med (2013) 5, 1642–1661
6. Gene therapy research in Asia, Gene Therapy (2017) 24, 572–577
7. Cell & Gene Therapy Market Survey 2019, Oxford Global
8. Oligonucleotide synthesis market – Global Forecast to 2022, MARKETANDMARKETS
9. Ryan Cross, C&EN, November 12, 2018, Volume 96, Issue 45
10. Lundin et al., Human Gene Therapy, Volume 26 Number 8, pp 475
11. Global Oligonucleotide Testing Market Size and Forecast 2014-24, Hexa Research

# A global perspective



**Stephen Lock**  
Business Development &  
Marketing Manager, CE, EMEA



**Yi Zhang**  
Market Development  
Manager-Pharma/CRO, Americas



**Ferran Sanchez**  
Market Development  
Manager, Pharma,  
EMEA



**Hans Dewald**  
Senior Marketing  
Manager, Americas



**Wenhai Jin**  
Director, Market Development,  
Life Sciences, China



**Zhe Zhou**  
Market Development  
Manager, China



**Peng Zhao**  
Marketing Development  
Manager, CE & Biopharma,  
China



**Kazunari Kessoku,**  
Market Development  
Manager, Japan



**Kaoru Karasawa**  
Application Support Senior  
Specialist, Japan



**Seo-Bong Chang**  
KOSEA CE  
Technical Champion



**Jason Neo**  
Director, Marketing & Field  
Applications Support, KOSEA



**Shigeo Hatsusaka**  
Market Development Manager,  
CE & Biopharma, Japan

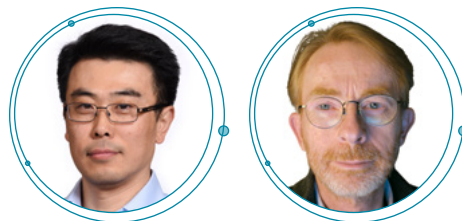


# Regional landscape

## North America

**Yi Zhang**, Market Development Manager-Pharma/CRO, Americas

**Hans Dewald**, Senior Marketing Manager, Americas



There is no doubt that the gene therapy market is one of the fastest growing sectors within the global pharmaceutical market today. North America is fueling much of this growth with about 50% of all new developments, including clinical trials, happening in the United States and is projected to grow at a CAGR of 41%. The high level of public and private funding in gene therapy within the US supports this growth. Likewise, there are over 100 clinical trials for oligonucleotide therapeutics in the US with five new oligonucleotide therapies approved by the FDA in the last two years. Similar to the rise of monoclonal antibody therapeutics over the last 20 years, much of the initial interest and activity in oligonucleotides and gene therapies started in America then quickly circled the globe with market acceleration following in Europe, Japan, and the rest of the world.

With respect to viral vector formats, typically AAV and lentivirus are the two major transport vehicles being

used for classical gene therapeutics in America. The primary driver around the use of AAV is cost. Because AAV only requires 2 or 3 plasmids, versus 4 with lentivirus, small companies or nascent programs have gravitated to AAV as the most economical solution. With respect to oligonucleotide therapeutics, the current trends are to address stability challenges with highly chemically modified nucleoside backbones and delivery challenges with either nanoparticle formulations and/or covalent modifications that target specific cell receptors.

To bring gene therapies and oligonucleotide therapies to market faster, small biotech firms are now partnering with large pharmaceutical companies in order to gain access to resources for development and commercialization. But ultimately, every company and country needs better solutions to these common analytical challenges in order to ensure timely access to life-saving gene therapy medicines for patients.



## EMEAI (Europe, Middle East, Africa and India)

**Stephen Lock**, Business Development & Marketing Manager, CE, EMEAI

**Ferran Sanchez**, Market Development Manager, Pharma, EMEAI



The gene therapy market in EMEAI, particularly in Europe, is following right on the heels of North America. Many of the larger companies within Europe have footprints in the US, so there are consistencies in programs and processes that are highly integrated between the two continents.

Europe holds roughly 30% of the gene therapy market with a 38% CAGR. At the country level, the UK, France, Germany and Switzerland are leading the way with their investments in the development of gene therapeutics, the creation of gene therapy networks and associations, such as the European Society of Gene and Cell Therapy (ESGCT). Other countries that are starting to have a substantial interest in gene therapies include Italy, Spain, Belgium, Hungary, Russia.

Several small company start-ups have been created based on research discoveries from universities and funding from venture capital firms or large pharma. As a result, centers or ‘hubs’ are emerging in cities around traditional biopharmaceutical companies or in university towns such as Cambridge in the UK.

In the European Union, the National Health Services enables patient care to be offered across the region. There is a fluidity within the eurozone where patients have the option to travel to other countries for treatment. As a result, a national health system tourism has emerged. For gene therapies, the consensus is largely the same in that therapies will be offered consistently around the EU.



## China

**Wenhai Jin**, Director, Market Development,  
Life Sciences China

**Zhe Zhou**, Market Development Manager, China

**Peng Zhao**, Marketing Development Manager,  
CE & Biopharma, China



The gene therapy market in China continues to evolve. In China, there are perhaps four main forces that are helping to shape the market evolution. The first influence is a result of the demographic explosion within China. With such a large population, the need for new and effective drug treatments is in high demand and expected to only increase further. The gene therapy market in China is now growing at a compound annual growth rate (CAGR) of more than 25% year over year. China was the first country to commercialize gene therapies in 2003 and the demand for new gene and cell therapies will only grow stronger as we move forward.

Unfortunately, product launches since 2012 have been very slow. Part of this slowdown can be attributed to the second factor shaping the China market, which is the lack of a clear policy around the development and manufacturing requirements of gene therapeutics. Indeed, two competing regulatory tracks, also known as the “dual track management” process, has slowed commercialization as government agencies try to cope with the complex questions around how to safely bring gene and cell therapies to market. This dual track system is greatly influenced by the policies of the Chinese government and is an evolving landscape. As a result, there are currently no large Chinese pharmaceutical companies with a singular focus on gene therapeutics. Instead, many smaller companies and divisions have emerged with collaborations forming between many Chinese and foreign institutions.

The uncertain policy terrain within China has led to a slowdown in new innovations in gene therapy research – the third factor influencing overall market dynamics.

Discovery and development programs within China have stalled compared to other regions of the world. In the past, vector research dominated gene therapy. Now, collaborations with foreign companies around engineered chimeric antigen receptors on T-cells (CAR-T cell therapies) account for 50% of research and clinical trials. Because the most efficient way to transfect T-cells is using lentivirus, lentiviral products are more popular than AAV for cancer solutions in China. Additionally, because of the ongoing policy issues in the gene therapy space, lentivirus and plasmid products are imported into China rather than produced locally.

Finally, drug prices and development costs play a major role in how the market evolves. Because of new policies recently introduced around the promotion and pricing of generic drugs, China pharma companies are increasingly looking towards gene therapy and other modalities to gain a competitive and financial advantage. In addition to the cell based therapy collaborations previously mentioned, several large pharma companies are now adding oligo therapies and RNAi therapies into their pipeline. But ultimately, the cost of any new drug treatment will play a major role in market adoption within China. Any cost savings that can be accomplished along the path towards commercialization are beneficial as they have the potential to be extended to the price of the end product. Now more than ever, there is a clear need for analytical tools and methods that can provide fast and comprehensive analysis of viral vectors, plasmids, oligos and other gene and cell therapy modalities.



## Japan

**Kazunari Kessoku**, Market Development Manager

**Kaoru Karasawa**, Application Support Senior Specialist, Japan

**Shigeo Hatsusaka**, Market Development Manager, CE & Biopharma

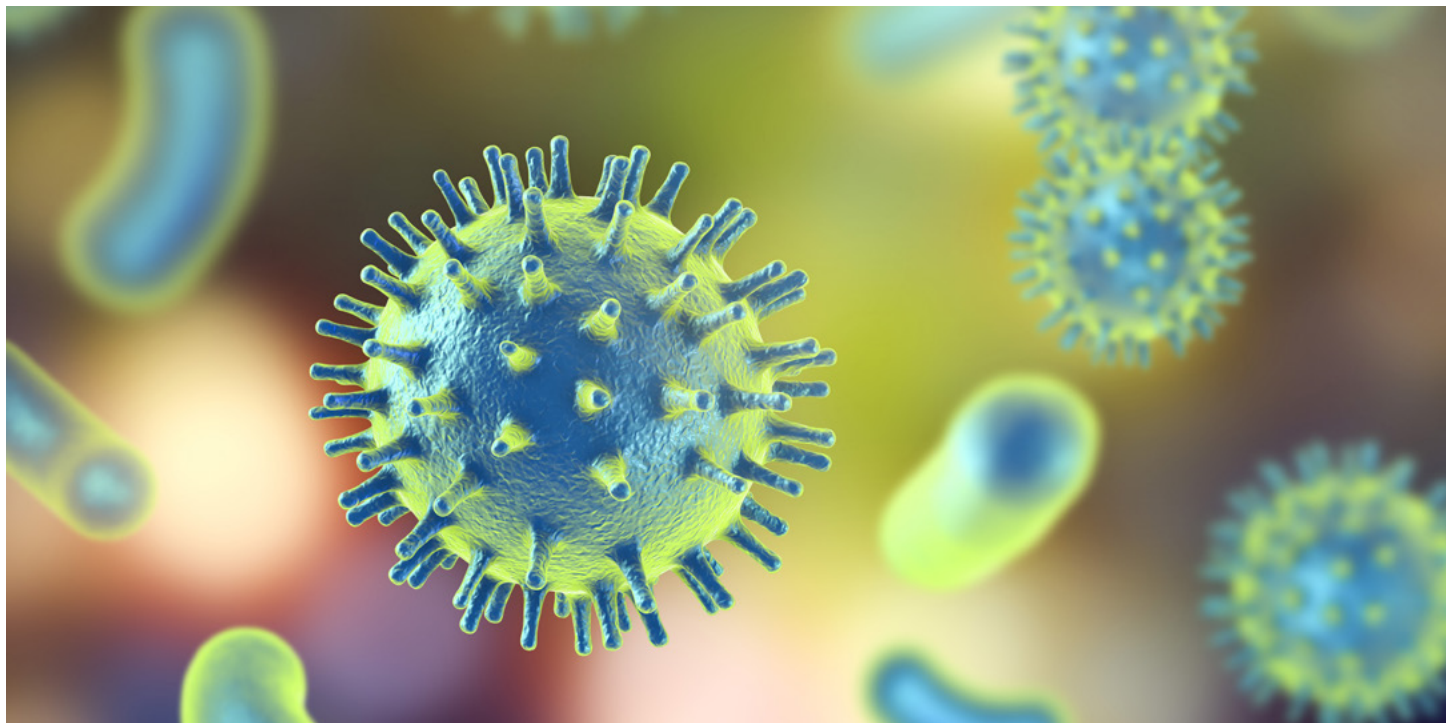


Although the rise of the global gene therapy market began over 30 years ago, it wasn't until recently that the market started to grow more rapidly. In Japan, the growth has not been as fast compared to Europe and America. However, significant research and development is progressing with approximately 70% in the preclinical research phase and 30% in the clinical trial phase.

In vivo gene therapy is being developed for various diseases in areas such as oncology, ophthalmology and inherited or rare diseases. On the other hand, ex vivo gene therapies are concentrated in the oncology area with approximately 70% of development in this area.

These results indicate that in vivo gene therapy is more diverse as a therapeutic agent than ex vivo gene therapy and might be used in the treatment of a larger set of diseases in the future. As for viral vector research and development, the main players are retroviruses, adenoviruses, plasmids, lentiviruses and AAV vectors as well as CAR-T.

Although no product has been approved yet in Japan, currently more than 60 clinical trials from Japanese research programs have been approved with some of these trials being conducted in Japan. With the global gene therapy market projected to have explosive growth over the next 10 years, a similar trend is expected in Japan, so expectations for the gene therapy market are very high.



## Korea and South East Asia

**Jason Neo**, Director, Marketing & Field Applications Support, KOSEA

**Seo-Bong Chang**, KOSEA CE Technical Champion



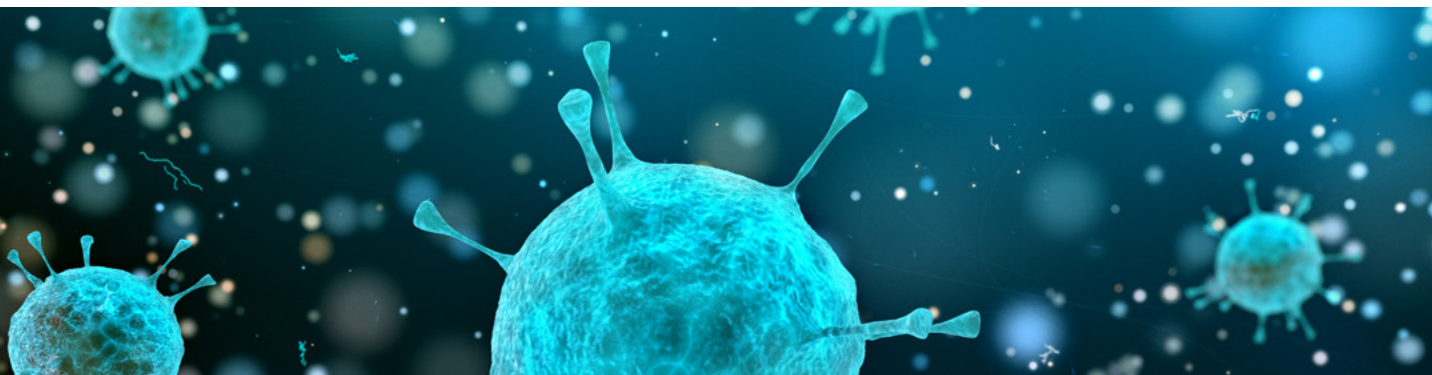
The gene therapy market in Korea and South East Asia is still very young. Within the region, the number of new gene therapies currently within the regulatory approval process is 18 in Korea, 17 in Australia, 8 in SE Asia, 6 in Taiwan and 3 in Hong Kong. Most of the companies developing gene therapies are small biopharmas and start-up enterprises. Government institutions and large pharmaceutical firms are also starting to enter the market with dedicated gene therapy programs. Nonetheless, the market is still in its early stages compared to the rest of the world and with different maturity levels depending upon the country.

For example, gene therapy is a very popular topic at virtually all biologics meetings in Australia. Many institutions want to enter the gene therapeutics arena and funding is becoming available. However, regulations are still being developed and scientists are looking for direction. Scientists have technical and analytical challenges that need solutions. Many companies do their manufacturing overseas and then bring the products back locally to introduce them into their viruses. As a result, much of the interest lies in evaluating the infectivity of the virus and the sample loading. Thus, analytical methods for full/empty capsid

analysis are in high demand in addition to methods for purity analysis of the final product.

Conversely, in Korea, the gene therapy market is probably still 2-3 years away from true acceleration. There is some interest around AAV analysis and QC methods, but most of Korea is focused on complex biologics molecules. Because the gene therapy market is still so young, there is not much regulation yet. However it is generally believed that the requirements will ultimately be aligned with the development process for biologics.

There are clearly differences starting to emerge within certain countries around choice of viral vector. For example, there is a lot of emphasis in using retrovirus in Australia, but AAV seems to be the dominant preference in Singapore and Korea. Besides using CE systems for analysis, there is also interest in using mass spectrometry based solutions. Once such solution is quantifying the particles themselves using a quick method to make sure the right capsid component is created and purified. Ultimately, the availability of fast and accurate tools for the analysis of gene therapies in research and development will help to advance the gene therapy markets for all countries in Korea and Southeast Asia.



# Trusted precision to help you accelerate drug discovery and development



## Get fast, accurate results for gene therapies and oligonucleotides

Are you looking to bring safe, effective therapies to market sooner? A complete suite of instruments and solutions from SCIEX can help. Through the precision analytics of SCIEX CE and LC-MS/MS technology, get the answers you need to accelerate your drug development. Whether you are investigating ASOs, siRNAs, mRNAs, CRISPR, gene therapies or nucleic acids, SCIEX has your gene therapy and oligonucleotide research and development needs covered.

**Discover ways to accelerate your gene therapy and oligonucleotide development by getting the right answers the first time, through precise analytics you can trust.**

**For more information visit [sciex.com/GT-Solutions](http://sciex.com/GT-Solutions)**



The Power of Precision

# Modalities

## Viral Gene Therapy

Viral gene therapies are recombinantly produced systems composed of a genetic construct consisting of essential viral genes, a therapeutic gene (transgene) to be introduced and associated promoters, all contained within a viral capsid. Although there are many types of viral vector systems, by far the 2 most commonly used are adeno associated virus (AAV) and lentivirus (LV). Viral gene therapies offer the capability of introducing extensive amounts of genetic instructions to a wide variety of target cell types with high efficiency to produce nearly any type of protein. Unfortunately, viral gene therapies are very difficult to produce in high yields with high purity. Viral vector biosynthesis typically requires multiple plasmid transfections to HEK293 cells with varying efficiencies and outcomes with each transfection step. Timely and accurate analysis of transfection steps and viral production will lower the cost as well as improve the safety of these therapeutics.



## Oligonucleotides

Therapeutic oligonucleotides are chemically synthesized strands of nucleotides in the range of 15-30 bases. Therapeutic oligonucleotides are broadly broken down into 2 classes: antisense oligonucleotides (ASOs) and small interfering RNAs (siRNAs). Whether ASO or siRNA, these oligonucleotide sequences are highly chemically modified to increase both plasma and intracellular stability. Both classes selectively target the mRNA transcript of the protein to be blocked, utilizing either RNaseH for ASOs or RISC for siRNAs or the pre-mRNA transcript of the protein to be modulated via differential splicing. Therapeutic oligonucleotides are difficult to produce in high yields with high purity, and they are difficult to formulate for efficient delivery and stability. Furthermore, oligonucleotides have a degree of regulatory uncertainty since their biophysical characteristics straddle the line between small molecules and biologics.





## CRISPR

CRISPR (clustered regularly interspaced short palindromic repeats) sequences and Cas (CRISPR associated) proteins are a biological system for permanent genome editing. The CRISPR/Cas system consists of a recombinant Cas (typically Cas9 or CPF1) enzyme that catalyzes double stranded DNA cleavages with extraordinary specificity based on hybridization of the synthetic CRISPR targeting sequence (sgRNA) against the target genomic DNA. Furthermore, HDR edited or corrected DNA sequences can be inserted into the genome. As previously mentioned, the CRISPR/Cas system is a combination of both synthetic and recombinant components. Each component needs to be independently manufactured and then formulated together. An emerging approach is to combine both guide RNA sequence and Cas mRNA protein sequence into a single plasmid. Nonetheless, both the single plasmid and plasmid/RNA combination requires either formulation into lipid nanoparticles or packaging into a viral vector for delivery. An additional challenge demonstrated by investigational CRISPR/Cas therapies is the need to control dose and/or nuclear residence time to carefully balance efficacy with off-target effects.

# Technical note summaries

## Capsid protein analysis

***One of the most common vehicles for delivering gene therapy is Adeno-associated virus or AAV. The virus is composed of a single-stranded DNA surrounded by a shell of proteins called capsid. Three main proteins make up the viral capsid: VP1 (87kD), VP2 (73kD) and VP3 (61kD). In order to ensure the safety and efficacy of AAV products, purity analysis of capsid proteins is extremely important.***

### Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins using CE-SDS Method

This study illustrates the separation, characterization and quantification of capsid proteins using the PA 800 Plus System with greater confidence, precision and accuracy compared to traditional methods such as SDS-PAGE. The PA 800 Plus System interfaced with a PDA detector provides analysis to separate, characterize and quantify capsid proteins with higher resolution, better quantification capability and better reproducibility than traditional SDS-PAGE. Excellent linearity with  $R^2$  of 0.9991 was obtained from  $5 \times 10^{11}$  GC/mL to  $1 \times 10^{14}$  GC/mL for AAV samples with RSD less than 0.7% over 8 consecutive injections. In addition, the sample preparation was straightforward and less labor intensive than traditional SDS-PAGE.

### Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins using CE-LIF Technology

Because of the speed, reproducibility, resolution and ease of sample preparation, capillary electrophoresis sodium dodecyl sulfate (CE-SDS) using UV detection is an excellent approach for capsid purity analysis. However, for ultra-low-level AAV capsid purity analysis such as for in-process AAV product analysis, fluorescence detection can provide a higher degree of sensitivity.

Using the PA 800 Plus System, a capillary electrophoresis laser-induced fluorescence (CE-LIF) method was developed for purity assessment of AAV capsid proteins. The method shows excellent resolving power, good repeatability and good linearity for characterizing and quantifying AAV viral proteins. Additionally, it enables samples with concentrations as low as  $1 \times 10^{10}$  GC/mL to be analyzed with an easy and straightforward sample preparation protocol.

### Sensitive AAV Capsid Protein Impurity Analysis by CE Using Easy to Label Fluorescent Chromeo Dye P503

An easy-to-use two-step “denature and label” sample preparation method was combined with capillary electrophoresis sodium dodecyl sulfate (CE-SDS) and laser induced fluorescence detection (LIF) for trace level detection of AAV8 capsid proteins and impurities. Using the PA 800 Plus Pharmaceutical Analysis System and LIF detection, the assay enabled the very sensitive detection of AAV8 capsid proteins and impurities down to  $1 \times 10^9$  GC/mL, or 50 ng/mL range. Here samples are labeled using a pyrilium based fluorescent tag, Chromeo P503, that upon conjugation increases fluorescence quantum yield by 50% and undergoes a bathochromic shift of 100 nm. No buffer exchange or free-dye clean-up is required, further streamlining the sample preparation. Additionally, the CE-SDS MW chemistry kit was used with no modifications to the chemistry along with preassembled bare fused silica cartridges. Using a formulation buffer with low salt concentration and slightly higher pH allowed for high sensitivity due to favorable electrokinetic injection conditions and an advantageous labeling environment while still maintaining AAV8 stability.

---

## Nucleic acid sizing analysis

***Accurate sizing analysis of double and single stranded DNA and RNA is required in many areas of research, development and manufacturing of gene and cell therapies and other oligonucleotide based therapeutics. While traditional methods for analyzing nucleic acid-based molecules have relied on the use of slab gel electrophoresis, this approach can suffer from poor resolution, accuracy and reproducibility and limitations when analyzing longer length molecules or fragments. Alternatively, approaches using capillary electrophoresis can greatly improve upon these specifications, even for longer strands.***

### Accurate Analysis of Double Stranded DNA over an Extended Size Range: Capillary Electrophoresis - Laser Induced Fluorescence Detection (CE - LIF)

This study demonstrates dsDNA fragment analysis over an extended molecular size range from 100 base pairs (bp) to 15 kilo base pairs (kb) using capillary electrophoresis laser-induced fluorescence (CE-LIF) on the PA 800 Plus System in combination with a complete set of reagents (capillary, standards and dye). This tech note highlights different methods, all with good sensitivity, dynamic range and reproducibility. Users have the option to choose different capillary lengths, DNA standards, sample injection and separation conditions depending upon their analysis needs for different size ranges.

A key factor that enables greater accuracy in size determination is the ability to co-inject standards with the samples. As an added benefit, the shelf life of 15 months for the system kits and components is substantially longer than microchip based reagent kits that tend to have shelf lives of only 4 to 6 months.

### Method Evaluation for RNA Purity Analysis Using CE-LIF Technology

Purity and quality analysis of RNA is critical to ensure the safety and efficacy of gene therapy products and intermediates. Using capillary gel electrophoresis on the PA 800 Plus system equipped with laser-induced fluorescence detection, a method was optimized for the evaluation of RNA purity and quality. Different separation voltages and temperatures were compared to find the optimal resolution and peak shape. Additionally, other parameters such as 2% PVP in the separation buffer, pressure injection vs. electrokinetic injection and capillary length were also evaluated for their effects on analysis time, resolution, peak shape and stability of RNA sample solutions.

The final optimized method showed that baseline resolution could be achieved for RNA ladders from 0.2 kb to 6.5 kb with excellent linearity, repeatability and sensitivity, with 0.33 ng/mL LOQ. Thus, the calibration curve that was generated from the RNA ladders can be used to estimate the size of unknown sample peaks.

### RNA Analysis for CRISPR by capillary electrophoresis with laser-induced fluorescence detection

Using the PA 800 Plus System and the eCAP ssDNA 100-R kit, a capillary electrophoresis laser-induced fluorescence (CE-LIF) method was developed for purity analysis of both sgRNA and Cas9 mRNA using the same gel matrix. The method was able to analyze small and large RNA in one sequence with single base resolution down to 15 nucleotides in length. Using a diluted gel matrix, RNA size estimation could be performed in the range of 0.2 kb to 10 kb which allowed detection and analysis of sequences from RNA degradation.

The method provided excellent repeatability and the single base resolution down to 15 nt provides confidence in analyzing n-1 species in short oligos such as crRNA and tracrRNA. Additionally, the eCAP ssDNA 100-R kit supplies pre-formulated gels. The entire method not only simplifies the analysis process, but also saves time and resources.

### Highly Precise DNA Sizing on the GeXP Genetic Analysis System

In this study, the GenomeLab GeXP™ Genetic Analysis System was evaluated for its ability to precisely size microsatellite amplification products. Four different products with sizes of 196, 277, 281 and 403 bp were accurately sized with a standard deviation of less than 0.27 bases. This high level of precision was achieved regardless of the choice of capillary within the array, gel lot, instrument and run. Because of this high precision, proper identification of alleles that differ by only 1 nucleotide can be accomplished for sizes up to 400 bases.

While, the use of internal size standards enables a higher level of precision that far exceeds the visual comparison of adjacent lanes on a slab gel, much of the outstanding reproducibility that was observed here can be attributed to additional factors. For example, the consistency in the manufacturing of the gels, the coating of the capillaries, the uniformity of capillary temperature and run voltage and the ability of the software to accurately identify peak centers all contribute to the high level of precision that can be achieved.

---

## Plasmid analysis

***Plasmid DNA is used as a gene delivery vehicle and in the manufacture of AAV and lentiviral products for gene therapies. In order to ensure the safety and efficacy of plasmid DNA, the determination of the relative amounts of different plasmid isoforms (e.g., linear, supercoiled, etc.) is an important practice when working with plasmid products or intermediates.***

### Simple, Fast and Robust Method for Plasmid Purity Testing and Degradation Monitoring: Capillary Electrophoresis-Laser Induced Fluorescence Detection (CE-LIF)

This study demonstrates the sensitive quantitative analysis of 5 kb and 7 kb plasmid molecular isoforms with the PA 800 Plus equipped with laser-induced fluorescence in a highly automated workflow. An important aspect of the workflow involves dilution of the eCAP dsDNA 1000 gel with 1X TBE buffer in order to provide highly resolved plasmid isoforms in a short period of time. In less than 12 minutes, baseline separation of supercoiled, linear and open circular plasmid isoforms is achieved. The method is capable of monitoring low-level plasmid degradants with 0.23% RSD for migration time and 0.61% RSD for % area. Analysis of plasmids larger than 7 kb or significantly smaller than 5 kb can be achieved with more or less dilution of the dsDNA 1000 gel.

Compared with slab gel methods that use ethidium bromide, the LIFluor EnhanCE fluorescent stain used in this method is safer. Additionally, less waste is generated, sample handling is minimized, sample loading is automated, greater sensitivity and reproducibility are observed and more accurate quantification is achieved.



### A sensitive and robust plasmid analysis method by Capillary Electrophoresis-Laser Induced Fluorescence

This study demonstrates highly robust and accurate characterization of plasmids using the PA 800 Plus Pharmaceutical Analysis System. First, resolution is maximized by investigating different dilution factors of the gel matrix on sieving performance. Next, pressure injection versus electrokinetic injection methods are evaluated for their effects on resolution. And finally, four different dyes are compared in order to assess their impact on overall separation and sensitivity.

After optimization of the assay, 21 consecutive injections of a 10 kb plasmid are run to evaluate overall robustness and reproducibility. Baseline resolution of different topological isoforms is achieved with consistent peak profiles for all 21 consecutive runs and % RSD below 8%.

Finally, experiments with stressed plasmid samples show the detection of changes in the relative quantities of different isoforms over the course of 12 weeks. Compared with an agarose gel method, the CE-LIF method improves resolution and detection as well as provides better quantitative analysis.

## Selection of aptamers

*Aptamers are short single stranded DNA or RNA sequences with unique structures. In the SELEX method (Systematic Evolution of Ligands by EXponential enrichment) aptamers with high affinity and selectivity for a target molecule are isolated from a random sequence nucleic acid pool, amplified and then incubated again in an iterative process in order to selectively purify aptamers with strong binding affinities.*

### CE-SELEX: Rapid Aptamer Selection Using Capillary Electrophoresis

This study illustrates the use of the P/ACE™ MDQ Plus Capillary Electrophoresis System for CE-SELEX selection of aptamers compared with conventional SELEX. For example, the aptamer selected for Immunoglobulin E (IgE) obtained by conventional SELEX required 15 selection rounds with negative selections, compared to 4 rounds without negative selection for the aptamer selected using CE-SELEX. Additionally, Human IgE aptamers from 2-4 rounds of CE-SELEX had at least 50-fold selectivity over mouse IgE and 500-fold selectivity over human IgG. The CE-SELEX method easily supports limited samples with target concentrations as low as 1 pM and volumes as low as 5 µl. Furthermore, compared with conventional SELEX, CE-SELEX increases the number and types of viable targets, allows more flexibility in developing an optimized aptamer sequence, and is compatible with many non-natural nucleic acid libraries and modifications that cause issues for other SELEX techniques.

---

## Nucleic acid sequencing

***Nucleic acid sequencing is used extensively throughout the biological sciences and is an important tool for gene therapy development. Capillary based electrophoresis systems are now routinely used for sequencing because they can be highly automated and are easy to use.***

### Bacterial Artificial Chromosome (BAC) End Sequencing Analysis

This study demonstrates BAC (bacterial artificial chromosome) end sequencing for 6 different BAC clones using the SCIEX GeXP 8000 Genetic Analysis System and the GeXP DTCS kit or DTCS Quick-Start kit. All six clones were sequenced multiple times with relatively high signal levels, high-quality value scores and long read lengths. The overall average 98% cutoff base number from all six clones was 668 bp, whereas the overall average read length was 796 bp.

The method presented here enables rapid and efficient BAC end sequencing on a large scale with high degrees of accuracy, automation and throughput. The generic approach to the method allows the same methodology to be used for end sequencing of other large DNA constructs such as PAC, YAC and cosmid.

### Improved Sequencing of Plasmids on the GeXP 2000 By A Simple Template Preheating Procedure

This study describes how the use of heat can vastly improve the sequencing of long and GC-rich sequences on the GeXP 2000 System. Simply heating to 96°C for 1 minute prior to sequencing improves the sequence call for a GC-rich region from no bases to 704 bases with the first 637 bases being called with 98% accuracy. Applying the same heating method to the analysis of a large plasmid (>12 kb) shows that the treated sample called a total of 579 bases, whereas the untreated sample only called 461 bases.

The preheating procedure proves to be a simple, cost-effective and convenient method for greatly improving the automated sequencing of plasmid DNA templates.

### Method for Long Sequencing Read Lengths on the GeXP Genetic Analysis System

This study describes how the SCIEX GeXP Genetic Analysis System can produce sequence read lengths somewhat greater than 700 bases using the standard sequencing methods provided with the instrument. By modifying the separation voltage, duration, capillary temperature and injection time, SCIEX researchers were able to extend the 98% accuracy base read length cutoff to more than 900 bases, and in some cases, over 1000 bases.

Use of this new separation method increases the overall base sequence yield from a single reaction. Thus, the amount of sequence information that can be obtained from a single reaction is maximized.

---

## SNP analysis and genotyping

***Single nucleotide polymorphisms (SNPs) comprise approximately 80% of all known polymorphisms and are major contributors to genetic variation. They play a large role in our susceptibility to certain diseases as well as our response to certain drug treatments. Screening for genetic variants at the general population can highlight variants that are associated with disease and can aid in approaches to personalized healthcare.***

### SNPs, Mutations and DNA Sequence Variation Analysis using the GenomeLab™ SNPStart Kit

This study details how to use the single base primer extension (SBE) method with the GeXP 8000 Series Genetic Analysis System and a SNP Primer Extension Kit created specifically for SBE with all of the necessary reagents including fluorescent dyes, enzyme and buffer. The kit comes as a ready-to-use single tube reagent that is based on a simple primer extension technology with no additional secondary steps involved. The kit was used to demonstrate multiple levels of multiplexing capability using SNP panels from one to 10-plex. Multiplex Cystic Fibrosis (CF) mutation panels were then tested using patient DNA samples and all samples showed the correct genotype and mutation.

The SBE method using the GeXP 8000 system and SNP Primer Extension Kit is a fast, accurate and convenient method for SNP analysis and genotyping.

### Multiplex SNP Analysis: Screening Factor V R506Q (Leiden) Mutations

Using the GeXP 8000 Genetic Analysis System, GeXP DNA Size Standard 80 and GeXP SNP Primer Extension Kit, the single-base primer extension (SBE) method was used to screen for the Factor V mutation in the general population. Over 500 SNP reactions were performed to score this Factor V mutation, with the GeXP 8000 able to call the correct genotypes with virtually 100% accuracy. The SBE method as demonstrated here showed a high degree of throughput, automation and accuracy for validating and scoring the human Factor V mutation for venous thrombosis population. Additionally, the approach is generic enough that it can be easily adapted to validate or score thousands of other mutations associated with human disease.

### Multiplex Mutation Genotyping for Human Diseases: Breast Cancer and Familial Mediterranean Fever

This study demonstrates the use of the single base primer extension (SBE) method for the multiplexed analysis of BRCA1/2 mutations and MEFV mutations using the SCIEX GeXP 8000 Genetic Analysis System, GeXP DNA Size Standard 80 and GeXP SNP Primer Extension Kit. For the BRCA1/2 assay, 3 mutation loci were investigated. For the MEFV gene assay, 5 mutations were investigated with 2 of the mutations only one base pair apart. For both studies, SBE was able to detect mutations with a high degree of accuracy, distinguishing between individuals that are wild type, heterozygous or homozygous at all loci. As an added benefit, the SBE method required minimal optimization in both chemistry and software.

---

## Gene expression profiling

***Gene expression profiling can detect the activity of many genes at once, providing a more global view of a biological process or cellular function. By providing a relational view of a set of genes, with their interdependencies, modern gene expression profiling provides a powerful strategy for the discovery, optimization and development of novel drugs.***

### Quantitative Multiplexed Gene Expression: An Analysis of Sensitivity, Accuracy and Precision

The GenomeLab GeXP (Gene eXpression Profiler) Genetic Analysis System can detect up to 30 genes in a single reaction containing 5ng to 500ng of total RNA, or up to 240 samples per 24 hour period, producing 7,200 gene expression results for a 30-gene multiplexed assay. As demonstrated in the technical note, for a 10-point dilution series from 6.5ng to 250ng of Human Breast Cancer plex containing 24 genes, gene expression changes as small as 0.5-fold were confidently detected with relative accuracy > 96%. Excellent linearity and reproducibility were observed with average  $R^2$  of 0.9978 and %CV between 6.6-16.5 for the entire workflow (reverse transcription, PCR and capillary electrophoresis).

### Monitoring Multiple Pluripotency Biomarkers After Delivery of Reprogramming Factors to Human Somatic Cells

Gene expression profiles of human embryonic fibroblasts (HEF) and human adult retinal pigment epithelial (RPE) cells were examined following infection with recombinant lentivirus containing the cDNAs for five transcription factors. Using the GenomeLab™ GeXP Genetic Analysis System, a 25-gene multiplex study was performed to measure the expression profiles of the transcription factors over time, in addition to a number of other pluripotency markers. Expression of a complete set of endogenous, pluripotent marker genes was detected at day 30, strongly implying that these cells were de-differentiated.

Additionally, because retroviral vectors have the potential to activate proto-oncogenes and generate tumors, non-viral, particle-based technology consisting of single walled carbon nanotubes (SWNT) was also investigated for its ability to transfer a mixture of the five transcription factor cDNA plasmids.

The ability to observe minute changes in gene expression over the course of many days using the GeXP Genetic Analysis System was of great value in understanding the interplay of different factors and the time course for pluripotency induction from both lentiviral and SWNT delivery mechanisms.

### Confirming Gene Mutation by CRISPR-Cas9 at the Protein Level and Identifying Proteome-Wide Changes: Using SWATH® Acquisition on TripleTOF® Systems

Protein-level verification can be very advantageous for validating gene editing experiments, not only for inconclusive DNA- or RNA-based assays, but also for directly observing the effects of the gene editing event.

SWATH Acquisition using a TripleTOF® system for protein-level verification of gene editing experiments can provide a comprehensive view across the entire proteome, including the gene product in question. SWATH Acquisition was used for protein-level verification after CRISPR-Cas9 gene editing of the Major Vault Protein (MVP) in zebrafish. While mRNA verification was inconclusive, complete knock-out of the gene was confirmed with a number of additional proteins significantly up- and down-regulated. In total, ~3800 proteins were detected and quantified from each sample in a single run.



---

## Analysis of oligonucleotide therapeutics

*Because of its exquisite selectivity and sensitivity, mass spectrometry has found widespread use for many applications throughout the life sciences.*

*For the analysis of oligonucleotide therapies, mass spectrometry provides comprehensive characterization and quantification capabilities that support all stages of research, development and manufacturing.*

### High Resolution Analysis of Synthetic Oligonucleotides: Reverse Phase Ion Pairing LC-MS Analysis Using the X500B QTOF System

The detailed characterization of oligonucleotide therapeutics is extremely important in order to ensure the safety and efficacy of the final product. In this report, SCIEX researchers use the X500B QTOF system for purity assessment and sequence confirmation of oligos. The high mass accuracy and resolving capabilities of the instrument enable unambiguous confirmation of oligonucleotide mass. High quality chromatographic separation is accomplished using ion-pairing reverse phase chromatography with a generic gradient that can be applied to the analysis of a wide range of oligonucleotide sequences.

The method enables the unambiguous confirmation of sequences with minimal adduct formation or analysis induced modifications. Failure sequences, or impurities which are present in a sample, are highlighted in the software to quickly understand which samples have passed or failed. The resulting method provides a semi-automated solution for oligonucleotide characterization and QC.

### Extending the Lower Limits of Quantification of a Therapeutic Oligonucleotide Through Microflow LC-MS/MS Featuring SCIEX QTRAP® 6500+ LC-MS/MS System with OptiFlow™ Turbo V source and the SCIEX M5 MicroLC System

Accurate quantification of drugs in biological fluids is a necessary step for bringing potential therapies to market. In this technical note, a novel microflow strategy using the QTRAP® 6500+ LC-MS/MS System, OptiFlow® Turbo V Ion Source and M5 MicroLC System was used to increase the sensitivity, robustness and overall analytical performance for quantification of a phosphorothioate antisense oligonucleotide (fomivirsen).

Compared with an LLOQ of 5 ng/mL for standard flow at 300 uL/min, microflow at 6 uL/min provided a significant improvement with an LLOQ of 1 ng/mL. Improvements in peak area, peak intensity and signal to noise were also observed for microflow versus standard flow. Microflow also dramatically reduced the required mobile phase additives that inherently contribute to front end contamination of the mass spectrometer, saving on costs for reagents and buffers as well as decreasing downtime for routine cleaning.

### Quantification of Large Oligonucleotides using High Resolution MS/MS on the TripleTOF® 5600 System

Accurate quantification of large oligonucleotides presents several challenges to standard ELISA-based assays as they may not accurately distinguish large metabolites from the full-length oligonucleotide of interest. Using an MRMHR approach on the TripleTOF® 5600 System, SCIEX researchers developed a method for the quantitation of oligo therapeutics from biological matrices.

Oligonucleotides with masses of ~4500 Da were accurately quantitated achieving an LOQ of 0.05nM. Because the MRMHR method acquires a full scan MS/MS spectrum, specific fragment ions for quantification are chosen after acquisition. This greatly simplifies method development and provides an opportunity for high-throughput analysis of potential oligonucleotide-based therapeutics. Additionally, the high resolution MS/MS data that is acquired allows both quantitative and qualitative analysis to be performed in a single run.



# Capsid protein analysis

## Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins using CE-SDS Method

Tingting Li, Mukesh Malik, Handy Yowanto, Sahana Mollah  
SCIEX, Brea, CA

### Introduction

**Adeno-associated virus (AAV)** is one of the most widely used gene delivery vehicles for gene therapy because of its non-pathogenicity, low immunogenicity and different tropism to multiple cell types. It is made of a shell of protein called capsid encompassing a single-stranded DNA of about 4.8kb in size. The viral capsid is composed of 3 main proteins which are 87kD (VP1), 73kD (VP2) and 61kD (VP3) in size, respectively. Purity analysis of the AAV viral proteins is important for quality assurance and safety of AAV products.

Although SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) has been used for AAV capsid protein analysis in the industry, CE-SDS (Capillary Electrophoresis - Sodium Dodecyl Sulfate) method on the SCIEX PA800 Plus provides automated separation of proteins in the range of 10kD to 225kD with higher resolution, quantitation capability, better reproducibility and is less labor intensive than traditional SDS-PAGE.

This technical note demonstrates the capability of the CE-SDS method for purity analysis of AAV viral proteins with straight forward and easy sample preparation, excellent resolving power, good repeatability and linearity of absorbance response to sample concentration.

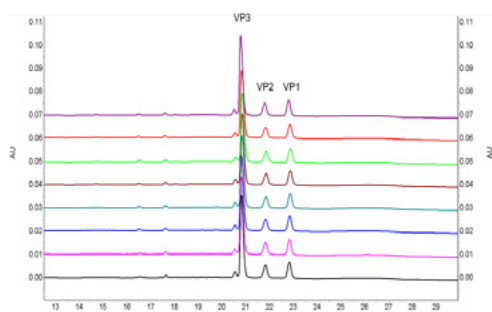


Figure 2. Eight consecutive injections of an AAV8 sample with estimated titer at  $8 \times 10^{13}$  GC/mL.



Figure 1. The PA800 Plus Biologics Analysis System (A), the SDS-MW Kit (B) and the EZ-CE Capillary Cartridge (C).

### Key Features

- Straightforward and easy sample preparation procedure
- Excellent Resolution of VP1, VP2 and VP3
- Excellent repeatability with RSD% of CPA% (Corrected Peak Area%) of viral proteins less than 0.7%
- Good Linearity of absorbance response to sample concentration with  $R^2 = 0.9991$

## Materials and Methods

### Materials

Sodium dodecyl sulfate (PN L4390-100G) and 2-mercaptoethanol (PN M3148-100ML) were obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.). Amicon Ultra-0.5 Centrifugal Filters with 30,000 NMWL were purchased from EMD Millipore (Billerica, MA, U.S.A.). The SDS-MW Analysis Kit (PN 390953) was from SCIEX (Framingham, MA, U.S.A.), which includes the SDS-MW gel buffer (proprietary formulation, pH 8, 0.2% SDS), acidic wash solution (0.1 N HCl), basic wash solution (0.1 N NaOH) and the SDS-MW sample buffer (100 mM Tris-HCl pH 9.0, 1% SDS). EZ-CE Capillary Cartridge (PN A55625, SCIEX, Framingham, MA, U.S.A) pre-assembled with bare fused-silica capillary (50  $\mu$ m I.D., 30 cm total length, 20 cm effective length) was used for separation. Universal vials (PN A62251), universal vial caps (PN A62250), PCR vials (PN 144709) and nanoVials (PN 5043467 from SCIEX (Framingham, MA, U.S.A.) were used for sample solution loading.

A PA800 Plus Pharmaceutical Analysis CE system (SCIEX, Framingham, MA, U.S.A.) equipped with a PDA detector and 32 Karat software were used for all the experiments. EZ-CE Capillary Cartridge (PN A55625, SCIEX, Framingham, MA, U.S.A.) pre-assembled with bare fused-silica capillary (50  $\mu$ m I.D., 30 cm total length, 20 cm effective length) was used for separation.

Data acquisition and analysis were performed using 32 Karat™ Software 10.

Packaged AAV2 of pAV-CMV-GFP with titer at  $2.24 \times 10^{13}$  GC/mL (titer as supplied by vendor) and packaged AAV8 of pAV-CMV-GFP with titer at  $3.99 \times 10^{13}$  GC/mL (titer as supplied by vendor) was purchased from Vigene Biosciences (Rockville, MD, U.S.A.). Both samples were kept in storage solution of PBS (Phosphate Buffered Saline, pH 7.5)/0.001% pluronic F68.

### Sample Preparation

**Sample Preparation Procedure.** 5  $\mu$ L of AAV8 sample solution with final salt concentration no more than 40mM was mixed with 5  $\mu$ L of 1% SDS and 1.5  $\mu$ L of 2-mercaptoethanol in a 0.65 mL micro-centrifuge tube and incubated at 50°C for 10min. Then, 90  $\mu$ L of DI water was added to the mixture. The diluted mixture was transferred to the sample vial or nanoVial for analysis on the PA800 Plus.

Buffer exchange is necessary if salt concentration in AAV sample is higher than 40mM.

**For Method Development and Optimization.** In this technical note, the method is developed and optimized using AAV8 samples

at  $1 \times 10^{13}$  GC/mL. Unless stated otherwise, 1.25  $\mu$ L of AAV8 sample at  $4 \times 10^{13}$  GC/mL and 3.75  $\mu$ L of deionized water was used in sample preparation procedure as 5  $\mu$ L of AAV8 sample solution at  $1 \times 10^{13}$  GC/mL.

All the titer values of AAV samples used in this technical note were provided by the vendor based on results from qPCR.

**For Method Evaluation.** To evaluate the capability of this method for analysis of AAV8 samples with lower titer, the AAV8 sample at  $4 \times 10^{13}$  GC/mL was diluted to  $1 \times 10^{12}$  GC/mL in the storage buffer provided by the vendor to represent the AAV8 samples at  $1 \times 10^{12}$  GC/mL.

Buffer exchange not only can be used to exchange the buffer with salt concentration lower than 40mM, it can be also used to concentrate the AAV sample for analysis. In this technical note, four times diluted storage buffer (salt concentration lower than 40mM) was used as elution buffer in buffer exchange procedure. The use of pluronic F68 is to minimize sticking of AAV to hydrophobic surfaces of plastics.<sup>1</sup>

Comparison of peak profiles of the AAV8 samples with and without buffer exchange was performed and discussed in the Result and Discussion section.

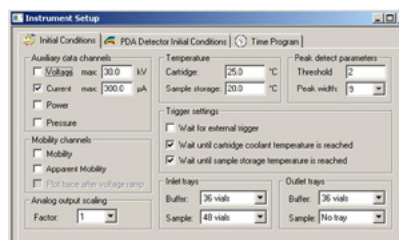


Figure 3. Initial Conditions.

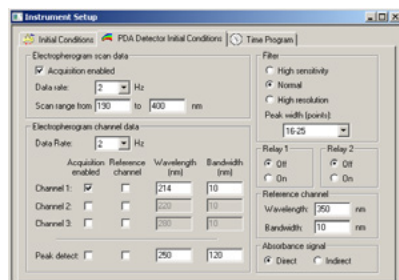


Figure 4. PDA Detector initial Conditions.





Other parameters such as starting sample volume for sample preparation, water plug for injections, sample injection duration, etc. were optimized for this procedure (Data not shown in this technote).

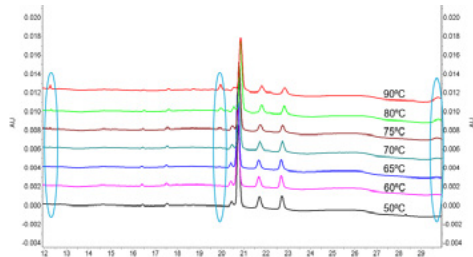


Figure 10. Optimization of incubation temperature for AAV8 sample preparation.

#### Method Evaluation

**AAV8 sample at  $1 \times 10^{13}$  GC/mL.** The method was developed and optimized using an AAV8 sample at  $1 \times 10^{13}$  GC/mL which is the nominal concentration from AAV manufacturing. Figure 11 illustrates the overlay of the 8 consecutive injections of this AAV8 sample. The VP3:VP2:VP1 ratio of the AAV8 sample lot tested is about 8:1:1.

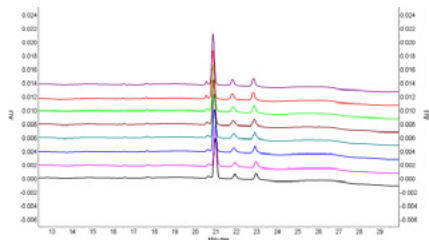


Figure 11. Consecutive injections of a  $1 \times 10^{13}$  GC/mL AAV8 sample.

**AAV8 sample buffer exchange to  $1 \times 10^{13}$  GC/mL.** Figure 12 shows the 6 consecutive injections of AAV8 sample buffer exchanged to  $1 \times 10^{13}$  GC/mL.

Comparison of Figure 11 and 12 showcase no difference in peak profile for AAV8 samples using different sample pretreatment methods (with or without buffer exchange).

**AAV8 sample at  $1 \times 10^{12}$  GC/mL.** This method was also evaluated for sample concentration as low as  $1 \times 10^{12}$  GC/mL. The AAV8 sample at  $1 \times 10^{12}$  GC/mL was prepared as described in Sample Preparation section.

The AAV8 samples with lower titer were buffer exchanged and concentrated before sample preparation procedure. Figure 13 shows 8 consecutive injections of the  $1 \times 10^{12}$  GC/mL AAV8

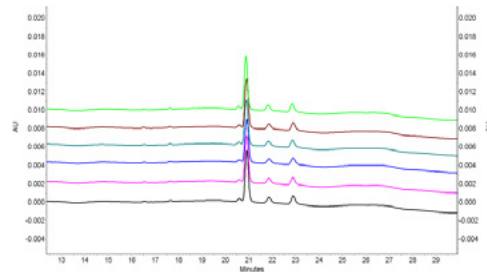


Figure 12. Consecutive injections of  $1 \times 10^{13}$  GC/mL AAV8 sample. AAV8 sample was buffer exchanged to  $1 \times 10^{13}$  GC/mL for sample pretreatment.

sample buffer-exchanged and concentrated to  $1 \times 10^{13}$  GC/mL (10 folds) for sample pretreatment.

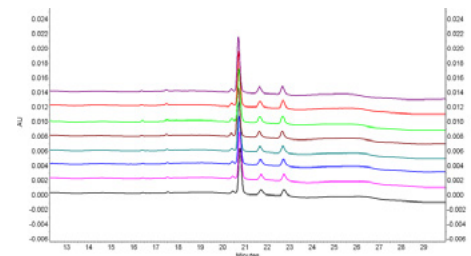


Figure 13. Consecutive injections of  $1 \times 10^{13}$  GC/mL AAV8. AAV8 sample was diluted to  $1 \times 10^{12}$  GC/mL with formulation buffer to represent samples with lower titer and then buffer exchanged to  $1 \times 10^{13}$  GC/mL for sample preparation.

Comparison of Figure 12 and 13 demonstrate no obvious difference in peak profile for AAV8 samples at different starting

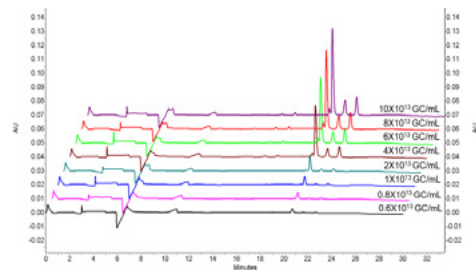


Figure 14. AAV8 sample was buffer exchanged to different concentrations (different folds).

concentration ( $1 \times 10^{12}$  GC/mL and  $4 \times 10^{13}$  GC/mL) using buffer exchange pretreatment.

**AAV8 samples concentrated to higher titer.** Figure 2 shows the overlay of 8 consecutive injections of AAV8 sample. The sample was buffer exchanged to  $8 \times 10^{13}$  GC/mL (2 folds higher) before sample preparation. The concentration after buffer exchange was measured as 0.7 mg/mL based on nanodrop reading at 280 nm.

**AAV8 samples at different titer.** A similar study was done by buffer exchanging the AAV8 sample to different titers (different folds) for method evaluation as shown in Figure 14. The titer values listed in the figure are a rough estimation from the folds of buffer exchange/concentration procedure.

**AAV2 samples (a different serotype).** The method was applied to an AAV2 sample. Figure 15 shows 8 consecutive injection of AAV2 sample buffer exchanged to  $0.5 \times 10^{13}$  GC/mL.

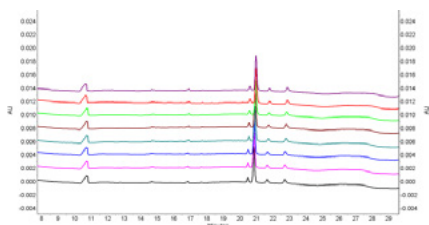


Figure 15. 8 Consecutive injections of  $0.5 \times 10^{13}$  GC/mL AAV2 sample.

**Repeatability.** All the consecutive injection studies using AAV2 and AAV8 samples at different concentration showcase good repeatability of peak profiles (Figure 2, Figure 11, Figure 12, Figure 13 and Figure 15). Figures of AAV samples at other concentrations are not shown in this technical note).

**Table 1. Repeatability of AAV2 and AAV8 at different concentration.**

Viral Proteins	RSD% of Corrected Peak Area						
	AAV8 _noBE_1	AAV8 _DBE_1	AAV8 _BE_8	AAV8 _BE_1	AAV8 _BE_4	AAV2 _noBE_0.5	AAV2 _BE_2
VP3	0.44	0.31	0.43	0.25	0.45	0.68	0.64
VP2	0.38	0.24	0.34	0.29	0.38	0.66	0.61
VP1	0.51	0.31	0.39	0.28	0.39	0.74	0.63

Table 1 demonstrates excellent repeatability of this method by evaluating the RSD% of corrected peak areas (CPA%) of 3 viral proteins of AAV2 and AAV8 at different titers and using different pretreatment methods. The calculation is based on 8 consecutive injections of each sample solution. All the RSD% of CPA% are

less than 0.7%. AAV8\_noBE\_1 is the  $1 \times 10^{13}$  GC/mL AAV8 sample without buffer exchange treatment. AAV8\_DBE\_1 is the  $1 \times 10^{12}$  GC/mL AAV8 sample buffer exchanged to  $1 \times 10^{13}$  GC/mL. AAV8\_BE\_8 is the AAV8 sample buffer exchanged to  $8 \times 10^{13}$  GC/mL. AAV8\_BE\_1 is the AAV8 sample buffer exchanged to  $1 \times 10^{13}$  GC/mL. AAV8\_BE\_4 is the AAV8 sample buffer exchanged to  $4 \times 10^{13}$  GC/mL. AAV2\_noBE\_0.5 is the  $0.5 \times 10^{13}$  GC/mL AAV2 sample without buffer exchange. AAV2\_BE\_2 is the AAV2 sample buffer exchanged to  $2 \times 10^{13}$  GC/mL.

**Linearity.** This method demonstrates excellent linearity of analyzing AAV8 samples from  $5 \times 10^{11}$  GC/mL to  $1 \times 10^{14}$  GC/mL by plotting absorbance response of VP3 to sample titers (Figure 16). The  $R^2$  is 0.9991.

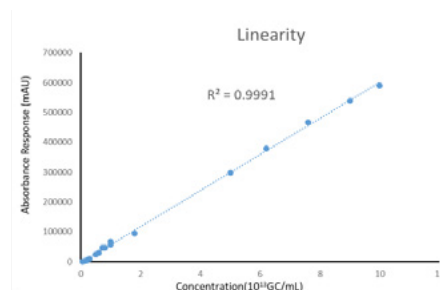


Figure 16. Linearity of Absorbance Response (mAU) to AAV8 Sample Concentration ( $10^{13}$  GC/mL).

## Conclusions

Presented in this technical note is a CE-SDS method for assessing the purity of Adeno Associated Virus (AAV) viral capsids. It uses a straight forward and easy sample preparation procedure and provides a good sensitivity for AAV samples which are of relatively lower concentration and of limited amount than monoclonal antibodies (mAbs). This method also demonstrates excellent resolution of the different viral proteins, good repeatability with less than 0.7% RSD% of CPA% of those viral proteins and good linearity of absorbance response vs sample concentration with  $R^2 = 0.9991$ .

## References

- <https://med.stanford.edu/gvvc/AAV.html>
- Zhang, C.; Meagher, M. M. *Anal. Chem.* **2017**, *89*, 3285–3292.
- Quirino, Joselito. (2015). Modern Injection Modes (Stacking) for CE. 10.1002/9783527678129.assep035.

For Research Use Only. Not for Use in Diagnostic Procedures.

## Drug Discovery and Development



AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

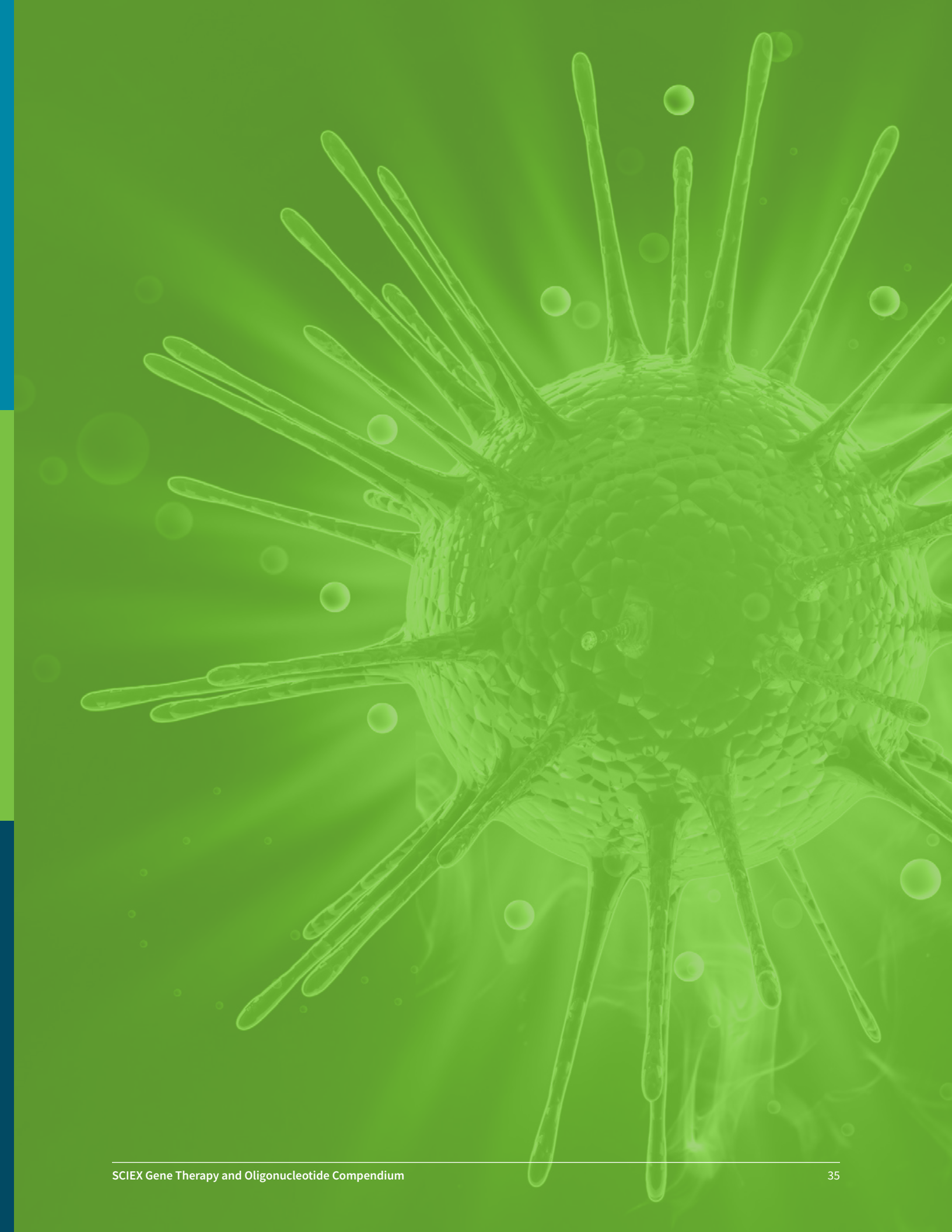
Document number: RUO-MKT-02-9761-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





## Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins using CE-LIF Technology

Tingting Li, Handy Yowanto, Sahana Mollah  
SCIEX, Brea, CA

### Introduction

**Adeno-associated virus (AAV)** is one of the most commonly used gene delivery vehicles in gene therapeutics. It is made of a shell of protein called capsid encompassing a single-stranded DNA of about 4.8kb in size. The viral capsid is composed of three main proteins which are 87kD (VP1), 73kD (VP2) and 61kD (VP3) in size, respectively. Purity analysis of the AAV viral proteins is important for quality assurance and safety of AAV products.<sup>1</sup>

CE-SDS (Capillary Electrophoresis Sodium Dodecyl Sulfate) technology is popular for AAV capsid protein analysis in the cell and gene therapy industry because of its automated separation of viral proteins with higher resolution, quantitation capability, better reproducibility and is less labor-intensive than traditional SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis). The CE-SDS method using UV or PDA detector and stacking injection technology could provide good results for AAV sample with titer greater than  $1 \times 10^{12}$  GC/mL.<sup>2,3,4</sup> However, for in-process AAV product analysis, higher sensitivity is required for purity analysis of AAV with concentration as low as  $1 \times 10^{10}$  GC/mL.

This technical note utilizes FQ (3-2-(furoyl quinoline-2-carboxaldehyde) dye for sample labeling and a laser-induced fluorescence (LIF) detector with 488nm solid-state laser and an emission filter of 600nm for sample detection to increase the sensitivity of AAV capsid purity analysis. It demonstrates the capability of this CE-SDS-LIF method for purity analysis of AAV viral proteins to as low as  $1 \times 10^{10}$  GC/mL with easy sample

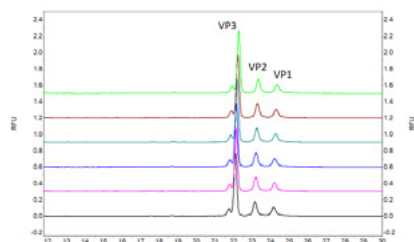


Figure 2. Six consecutive injections of an AAV8 sample with titer at  $1 \times 10^{13}$  GC/mL.



Figure 1. The PA800 Plus Biologics Analysis System equipped with a LIF detector (A), the SDS-MW Kit (B) and the EZ-CE Capillary Cartridge (C).

preparation, excellent resolving power, good repeatability and linearity of absorbance response to sample concentration.

### Key Features

- Ultra-high sensitivity AAV capsid purity analysis for AAV samples down to the in-process product analysis requirement level which is  $1 \times 10^{10}$  GC/mL
- Efficient labeling and simple sample preparation procedure in less than 1 hour
- Excellent baseline resolution of VP1, VP2 and VP3
- Good repeatability with RSD% of MT (Migration Time) less than 0.5% and CPA% (Corrected Peak Area%) of viral proteins less than 5% at  $1 \times 10^{12}$  GC/mL and  $1 \times 10^{13}$  GC/mL
- Good Linearity of absorbance response to sample concentration with  $R^2 = 0.9989$  from  $1 \times 10^{10}$  GC/mL to  $1.6 \times 10^{14}$  GC/mL

## Materials and Methods

### Materials

ATTO-TAG FQ and potassium cyanide which are included in ATTO-TAG™ FQ Amine-Derivatization Kit (PN A2334) were purchased from ThermoFisher Scientific (Waltham, MA, U.S.A.). Methanol (PN 1.06018.2500) was obtained from VWR (Radnor, PA, U.S.A.). Sodium dodecyl sulfate (SDS) (PN L4390-100G) and N-Ethylmaleimide (NEM) (PN E3876-5G) were obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.). Amicon Ultra-0.5 Centrifugal Filters with 10,000 NMWL (PN UFC501024) were purchased from EMD Millipore (Billerica, MA, U.S.A.). The SDS-MW Analysis Kit (PN 390953) (Figure 1B) was from SCIEX (Framingham, MA, U.S.A.), which includes the SDS-MW gel buffer (proprietary formulation, pH 8, 0.2% SDS), acidic wash solution (0.1 N HCl), basic wash solution (0.1 N NaOH) and the SDS-MW sample buffer (100 mM Tris-HCl pH 9.0, 1% SDS). EZ-CE Capillary Cartridge (PN A55625, SCIEX, Framingham, MA, U.S.A) pre-assembled with bare fused-silica capillary (50  $\mu$ m I.D., 30 cm total length, 20 cm effective length) was used for separation (Figure 1C). Universal vials (PN A62251), universal vial caps (PN A62250) and PCR vials (PN 144709) from SCIEX (Framingham, MA, U.S.A.) were used for sample solution loading.

A PA800 Plus Pharmaceutical Analysis CE system (SCIEX, Framingham, MA, U.S.A.) equipped with a laser-induced fluorescence (LIF) detector with a 600nm emission filter (Figure 1A) was used for all the experiments.

Data acquisition and analysis were performed using 32 Karat™ Software 10.

Packaged AAV2 of pAV-CMV-GFP with titer at  $2.24 \times 10^{13}$  GC/mL (titer as supplied by vendor), packaged AAV8 of pAV-CMV-GFP with titer at  $1.57 \times 10^{14}$  GC/mL (titer as supplied by vendor) and formulation buffer (Phosphate Buffered Saline, pH 7.5 with 0.001% pluronic F68) were purchased from Vigene Biosciences (Rockville, MD, U.S.A.). Both samples were kept in formulation buffer.

### Sample Preparation

**4% SDS in 150 mM NEM solution.** Weigh 93.75mg NEM and 0.2g SDS in 5mL DI water. Mix to dissolve.

**Labeling Solution.** The 20 mM FQ dye stock solution was prepared in methanol. This solution was further diluted in DDJ water to a final concentration of 2.5 mM and was used as the FQ dye working solution. The nucleophile reagent used was potassium cyanide at a concentration of 30 mM.

**1% SDS solution.** Dissolve 1g of SDS in 100mL DI water.

**Sample Preparation Procedure.** 10  $\mu$ L of AAV8 sample solution was mixed with 1.2  $\mu$ L of 4% SDS in 150mM NEM solution in a 0.65 mL micro-centrifuge tube and incubated at 70°C for 5 minutes. Then, the sample solution was mixed with 1.5  $\mu$ L of 2.5mM FQ dye working solution and 1  $\mu$ L of 30mM KCN solution and incubated at 70°C for 10 minutes. 28  $\mu$ L of 1% SDS was added to quench the labeling reaction. The sample solution was incubated at 70°C for 5 minutes. At last, 20  $\mu$ L of DI water was added to the mixture. The diluted mixture was transferred to the sample vial for analysis on the PA800 Plus.<sup>5</sup>

Buffer exchange or desalting to lower final salt concentration is an optional step before sample solution is loaded into sample vials to improve sensitivity of the method due to less competition from the salt during electrokinetic injection. The mixture of 5% of 1% SDS, 1.25% of formulation buffer in DI water was used as buffer exchange solution in this technical note.

### Instrument setup

The user guide of the SDS-MW Analysis kit was followed for instrument setup. Water plug was used in separation method for online sample concentration to improve the sensitivity of the method.<sup>2,3,4</sup>

The "Initial Conditions" and "LIF Detector Initial Conditions" were set up as indicated in Figure 3 and Figure 4, respectively. Same initial setup was used for conditioning, separation and shutdown methods.

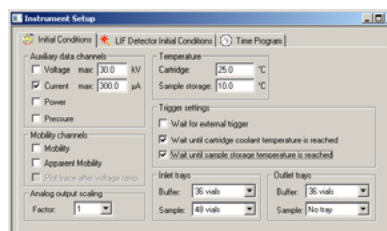


Figure 3. Initial Conditions.

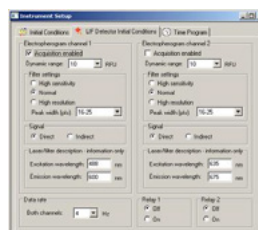


Figure 4. LIF Detector Initial Conditions.

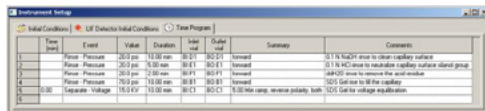


Figure 5. Time Program for Conditioning Method.

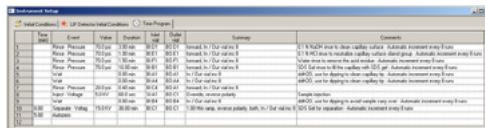


Figure 6. Time Program for Separation Method.

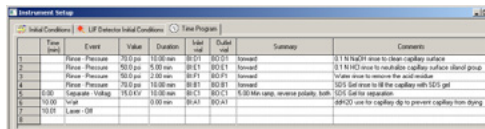


Figure 7. Time Program for Shutdown Method.

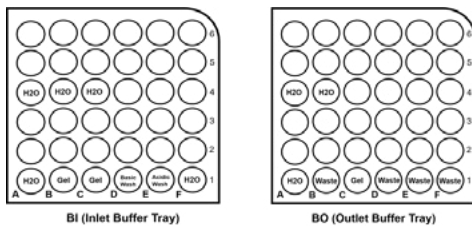


Figure 8. Buffer Tray Configuration.

The time program for conditioning method is illustrated in Figure 5. Figure 6 and Figure 7 are time programs for separation and shutdown method, respectively.

Figure 8 is the configuration for buffer tray setup.

## Result and Discussion

AAV8 and AAV2 sample at different titer used in this study were prepared by diluting the AAV8 sample at  $1.57 \times 10^{14}$  GC/mL and AAV2 sample at  $2.24 \times 10^{13}$  GC/mL by different factors in formulation buffer. If not stated otherwise, buffer exchange step was used for AAV samples at titer no more than  $1 \times 10^{12}$  GC/mL before loading into sample vials for PA800 plus analysis. No buffer exchange was done for samples in linearity study.

### PDA and LIF detector

AAV8 sample of  $1 \times 10^{13}$  GC/mL was used for comparison of capsid protein analysis results using PDA detector and FQ dye labeling

with LIF detector. Figure 9a is the electropherogram of  $1 \times 10^{13}$  GC/mL AAV8 capsid protein analysis using FQ dye labeling and LIF detector. Figure 9b is the electropherogram of  $1 \times 10^{13}$  GC/mL

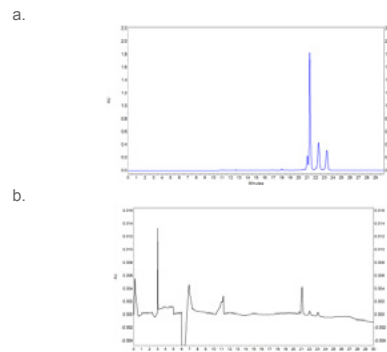


Figure 9. Comparison of capsid protein analysis using PDA detector and using FQ dye labeling with LIF detector. Figure 9a is the electropherogram of  $1 \times 10^{13}$  GC/mL AAV8 capsid protein analysis using FQ dye labeling and LIF detector. Figure 9b is the electropherogram of  $1 \times 10^{13}$  GC/mL AAV8 capsid protein analysis using PDA detector.

AAV8 capsid protein analysis using PDA detector (Sample preparation follow the procedure in Reference 2). The sample solution in Figure 9a was buffer exchanged to the same buffer as the one in Figure 9b before loading into sample vial. By comparing the electropherograms in Figure 9a and Figure 9b, the usage of FQ dye and LIF detector improves the sensitivity of CE-SDS method for AAV capsid purity analysis and provides a better and flatter baseline.

### Repeatability

This method was evaluated using AAV8 sample at different titers (Figures are not shown in this technical note).

Figure 2 shows the six consecutive injections of AAV8 sample at  $1 \times 10^{13}$  GC/mL. Figure 10 shows the six consecutive injections of AAV8 sample at  $1 \times 10^{10}$  GC/mL.

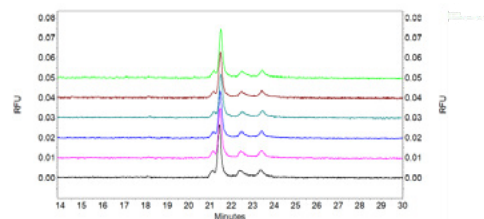


Figure 10. Six consecutive injections of an AAV8 sample with titer at  $1 \times 10^{10}$  GC/mL.

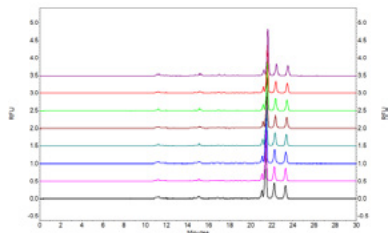


Figure 11. Eight consecutive injections of an AAV2 sample with titer at  $1 \times 10^{12}$  GC/mL.

This method was also evaluated using AAV serotype 2 sample at different titers. Figure 11 shows the eight consecutive injections of AAV2 sample at  $1 \times 10^{12}$  GC/mL.

**Table 1. Migration time repeatability of AAV2 and AAV8 at different concentration.**

Viral Proteins	RSD% of Migration Time					
	AAV8 _1X10 <sup>13</sup> GC/mL	AAV8 _1X10 <sup>12</sup> GC/mL	AAV8 _1X10 <sup>11</sup> GC/mL	AAV8 _1X10 <sup>10</sup> GC/mL	AAV2 _1X10 <sup>12</sup> GC/mL	AAV2 _1X10 <sup>10</sup> GC/mL
VP3	0.3	0.3	0.3	0.2	0.3	0.5
VP2	0.2	0.4	0.4	0.2	0.3	0.5
VP1	0.3	0.4	0.3	0.3	0.3	0.5

**Table 2. Corrected peak area% repeatability of AAV2 and AAV8 at different concentration.**

Viral Proteins	RSD% of Corrected Peak Area%					
	AAV8 _1X10 <sup>13</sup> GC/mL	AAV8 _1X10 <sup>12</sup> GC/mL	AAV8 _1X10 <sup>11</sup> GC/mL	AAV8 _1X10 <sup>10</sup> GC/mL	AAV2 _1X10 <sup>12</sup> GC/mL	AAV2 _1X10 <sup>10</sup> GC/mL
VP3	0.6	0.6	1.1	1.4	0.3	1.4
VP2	0.9	3.0	7.2	9.9	0.1	4.0
VP1	1.6	2.6	8.1	6.3	1.2	1.2

Table 1 and Table 2 demonstrates excellent repeatability of this method by evaluating the RSD% of migration time and RSD% of corrected peak areas% (CPA%) of three viral proteins of AAV2 and AAV8 at different titers. The calculation is based on six consecutive injections of each sample solution for AAV8 samples and 8 consecutive injections of each sample solution for AAV2 samples. AAV8\_1X10<sup>13</sup> is AAV8 sample with titer at  $1 \times 10^{13}$  GC/mL. AAV8\_1X10<sup>12</sup> is AAV8 sample with titer at  $1 \times 10^{12}$  GC/mL. AAV8\_1X10<sup>11</sup> is AAV8 sample with titer at  $1 \times 10^{11}$  GC/mL.

GC/mL. AAV8\_1X10<sup>10</sup> is AAV8 sample with titer at  $1 \times 10^{10}$  GC/mL. AAV2\_1X10<sup>12</sup> is AAV2 sample with titer at  $1 \times 10^{12}$  GC/mL. AAV2\_1X10<sup>10</sup> is AAV2 sample with titer at  $1 \times 10^{10}$  GC/mL. All the RSD% of MT% are no more than 0.5%. And the RSD% of CPA% are less than 10%.

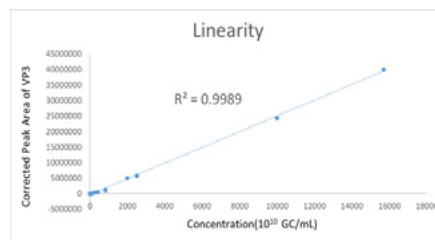


Figure 12. Linearity of corrected peak area to AAV 8 Sample Concentration ( $10^{10}$  GC/mL).

### Linearity

This method demonstrates excellent linearity of analyzing AAV8 samples from  $1 \times 10^{10}$  GC/mL to  $1.6 \times 10^{14}$  GC/mL by plotting peak area of VP3 to sample titers (Figure 12). The  $R^2$  is 0.9989.

### Conclusions

Presented in this technical note is a CE-SDS-LIF method for assessing the purity of Adeno Associated Virus (AAV) viral capsids utilizing FQ fluorescence dye labeling and LIF detector to provide an ultra-high sensitivity for in-process AAV product analysis with titer as low as  $1 \times 10^{10}$  GC/mL. The sample preparation procedure is straight forward and easy. This method also demonstrates excellent resolution of the different viral proteins, good repeatability and linearity of absorbance response vs. sample concentration with  $R^2 = 0.9989$ .

### References

- <https://med.stanford.edu/gvvc/AAV.html>
- SCIEX Technical Note: Purity Analysis of Adeno-Associated Virus(AAV) Capsid Proteins using CE-SDS Method
- Zhang, C.; Meagher, M. M. *Anal. Chem.* **2017**, *89*, 3285–3292.
- Quirino, Joselito. (2015). Modern Injection Modes (Stacking) for CE. 10.1002/9783527678129.assep035.
- SCIEX Technical Note: Using Fluorescent Labels to Increase the Sensitivity of IgG Purity and Heterogeneity Assay on PA 800 plus

For Research Use Only. Not for Use in Diagnostic Procedures.

## Drug Discovery and Development



AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-10086-A

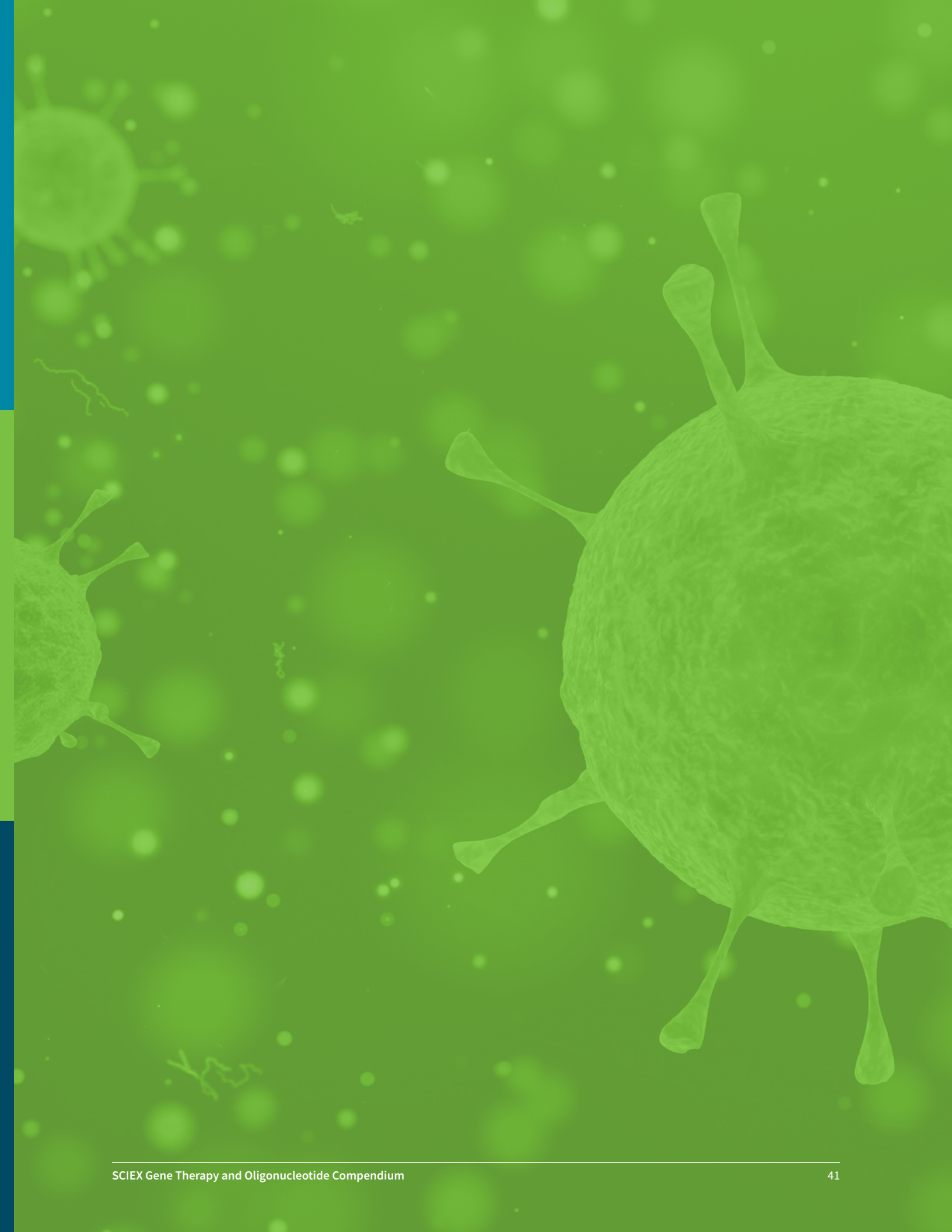


**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





## Sensitive AAV Capsid Protein Impurity Analysis by CE Using Easy to Label Fluorescent Chromeo Dye P503

Marcia Santos  
SCIEX, Brea, CA

### Introduction

AAV or Adeno associated virus (Figure 1) is a popular class of gene therapy delivery vehicle used mostly due to non-pathogenicity and high stability. Structurally, AAVs are composed of icosahedral proteic shell called Capsid or Viral Protein (VP) which contains a viral genome. So far, AAV exists in some 13 human and primate serotypes which in combination with the primary sequence differences mediate the AAV cell and tissue specificity. For example, AAV 8 or serotype 8 is efficient in transducing hepatocytes. This structure has the ability to carry up to 5 Kb of pay- load of single stranded DNA molecule. The capsid or viral proteins are translated from the same mRNA encoding overlapping sequences of three capsid proteins, VP1, VP2 and VP3 with approximately 87, 72 and 62 kDa, respectively resulting in a total of 60 monomers. In addition, in vivo, these viral proteins are synthesized roughly in a ratio of 1:1:10, respectively.<sup>1, 2</sup>



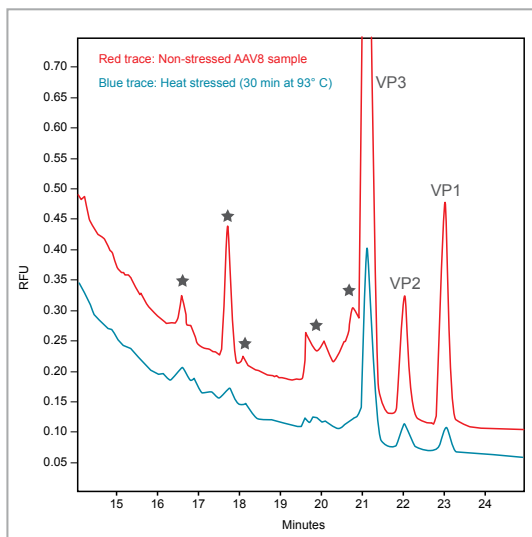
**Figure 1.** Pictorial representation of an adeno-associated virus.

As AAV take center stage in gene therapy treatment of many genetic conditions, reliable and quantitative assays are critical for the proper characterization of these molecules as well as the quantitation of impurities.

Typically, AAV concentration used in gene therapy is in the order of  $1 \times 10^{10}$  GC/mL (GC=genomic copies), which equates to 50 ng/mL and thus falls well below limit of detection or quantitation of any UV absorbance based assays.

In this technical note, an easy sample preparation and labeling scheme is described using commercially available fluorescent tag (Figure 3) which doesn't require buffer exchange nor dye clean-up for the low level detection of AAV8. Capillary Electrophoresis (CE) with Laser Induced Fluorescence (LIF) detection using a commercially SDS-MW chemistry kit and pre- assembled bare fused silica cartridge was successfully used to separate the impurities and characterize the AAV capsid proteins.

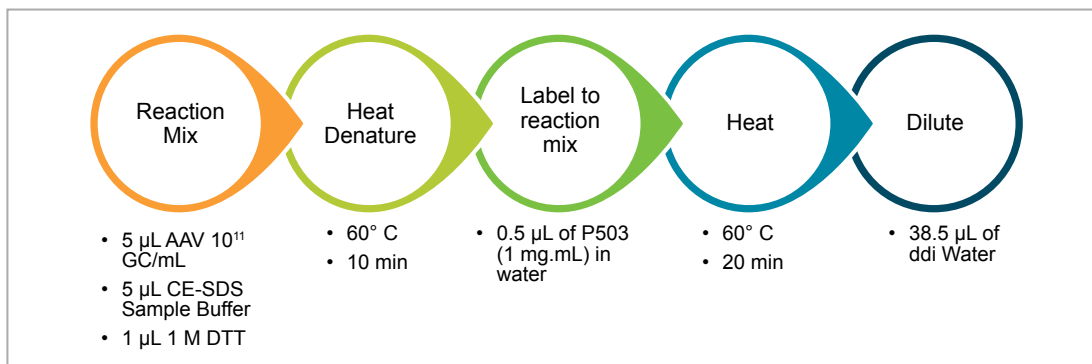
The fluorescent tag used in this work is a pyrium type of dye named Chromeo P503, which is reactive to primary amines. Chromeo P503 has very weak fluorescence as free dye; less than 1% quantum yield in solution. However, upon conjugation, not only does the fluorescence quantum yield rise to 50%, it also undergoes a substantial bathochromic shift of 100 nm.<sup>3</sup>



**Figure 2.** Overlay of Non-Stressed (red trace) and Heat-Stressed (blue trace) Both AAV8 Samples at  $1.1 \times 10^9$  GC/mL. Stars indicate regions on the electropherogram sensitive to heat treatment.

p1

[Download this technical note](#)



**Figure 3.** Simple Two Step "Denature and Label" Sample Preparation Scheme of AAV.

### Key Features

- Easy two step denature and label sample preparation.
- No buffer exchange or free-dye clean up required.
- Sensitive detection of impurities at the  $1 \times 10^9$  GC/mL or 5 ng/mL range, typical of therapeutic products, using Capillary Electrophoresis with LIF detection and Stacking techniques.
- Use CE-SDS MW kit with no modifications to the chemistry.

### Materials, Instrument and Methods

**Materials:** The CE-SDS MW kit (PN 390953) and LIF Performance Test Mix (PN: 726022) were from SCIEX, Framingham, MA. Packaged AAV8 of pAV-CMV-GFP with titer at  $1.10 \times 10^{13}$  GC/mL (titer as supplied by vendor) was purchased from Vigene Biosciences (Rockville, MD, U.S.A.). AAV formulation buffer (1X PBS with 0.001% Pluronic F68) was also from Vigene Biosciences. Chromeo P503 catalog # 15106, was purchased from Active Motif (Carlsbad, CA). Phosphate Buffered Saline Bioreagent Suitable for Cell Culture 10X, Sigma-Aldrich, PN P5493-1L was used as base for CE-SDS custom sample buffer. Sodium Dodecyl Sulfate, J. T. Baker, PN 4095-04.

**Sample storage:** Upon arrival, 5 µL aliquot of AAV8 sample were stored at  $-80^\circ\text{C}$  freezer to avoid multiple freeze-thaw cycles.

**Instrument and software:** A PA 800 Plus Pharmaceutical Analysis System equipped with LIF detector and solid-state laser with excitation wavelength at 488 nm were from SCIEX (Framingham, MA) and a 600 nm band pass emission filter from Edmund Optics (Barrington, NJ). Data acquisition and analysis

were performed using 32 Karat software™ V10. The separation method used in this work has already been described by Li.<sup>4</sup> Briefly, the CE separation method takes advantage of stacking technique by introducing a plug of water (20 psi/0.6 min) prior to the sample injections ( $-10$  kV/60 seconds).

**LIF Calibration:** To ensure consistent response of LIF detector throughout this study, the LIF detector was calibrated using LIF Calibration Wizard and Performance Test Mix (PN: 726022).

**Preparation of Chromeo P503 Working Solution:** A vial of Chromeo P503 dye comes in 1 mg of lyophilized powder. The lyophilized powder was reconstituted by adding 1 mL of methanol. Make 10 µL aliquots to prevent contamination due to over-handling. After reconstitution, the dye label can be stored at  $2-8^\circ\text{C}$  for six months according to the manufacturer's instructions.

**Sample Dilution Procedure Prior to Labeling:** One 5 µL aliquot of AAV8 sample was taken out of the freezer and diluted as follows: To prepare  $1.10 \times 10^{11}$  GC/mL: 1 microliter of AAV8  $1.10 \times 10^{13}$  GC/mL was added to a 99 µL of 50 mM phosphate buffer solution pH 8. To prepare  $1.10 \times 10^{10}$  GC/mL: 1 microliter of AAV8  $1.10 \times 10^{13}$  GC/mL was added to a 9 µL of 50 mM phosphate buffer solution pH 8.

**Sample Denaturing Procedure Prior to Labeling:** 5 µL of AAV8 diluted as in previous session were mixed with 5 µL of Tris sample buffer and 1 µL of 1M DTT. Both reaction mixes were briefly vortexed for proper mixing and heated to  $60^\circ\text{C}$  for 10 minutes. After, the reaction tubes were allowed to cool down to room temperature.

p2

Download this technical note

**Sample Labeling Procedure:** In this protocol, the least amount of sample possible is used, which may lead to challenges in pipetting very small volume. Scaling up the sample prep is possible and should be used if sample volume is not an issue. 0.5  $\mu\text{L}$  of 1 mg/mL of Chromeo P503 Labeling Working Solution was added to each reaction tube and briefly vortexed. Once again, both tubes were heated to 60° C for 20 minutes. Afterwards, both tubes were first allowed to cool to room temperature and then 38.5  $\mu\text{L}$  of DDI water was added to either reaction tube. After labeling and dilution, the final concentrations are AAV8  $1.1 \times 10^{10}$  and  $1.1 \times 10^9$  GC/mL respectively.

## Results and Discussions

**AAV Process and Product Related Impurities<sup>5</sup>:** Similar to a therapeutic monoclonal antibody production, there are stringent manufacturing requirements for the removal of product and process related impurities. Furthermore, sensitive analytical techniques for the detection and quantitation of impurities for product safety. This work is focused on the capsid impurities.

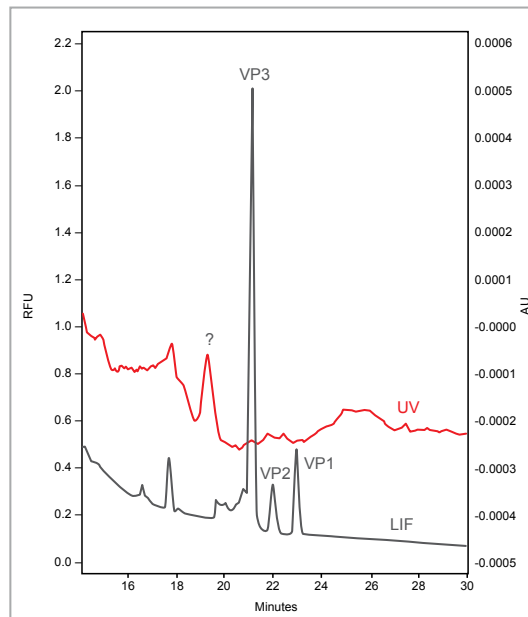
CE-SDS has been used as the gold standard for the characterization of therapeutic proteins from manufacturing to lot release. Combining the reproducibility and specificity typical of CE-SDS with sensitivity of LIF detection and easy to label sample preparation, AAV capsid protein purity can be easily determined.

Figure 4 shows a typical profile of P503 labeled AAV8 at the target concentration ( $10^{10}$  GC/mL) of clinical formulation compared to an absorbance UV. The labeling drastically improves sensitivity.

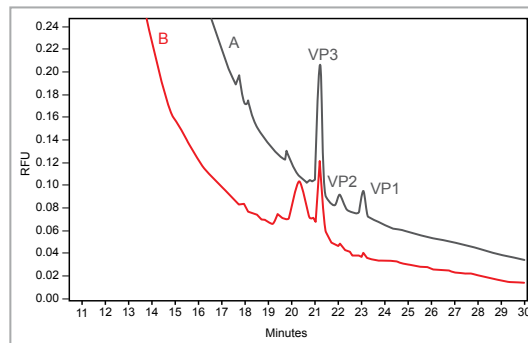
### Formulation buffer and the impact on assay sensitivity:

The type of injection typically used in CE-SDS assays is electrokinetic, where a low voltage is applied to drive charged ions into the capillary. However, there is an inherent bias towards highly charged small ions in electrokinetic injection. The combination of a large protein molecule of interest, such as capsid proteins, present in formulation buffer with high salt content will decrease the sensitivity of the assay, due to the dramatic difference in mobilities. Smaller ions move faster during electrokinetic injection than their larger protein counterparts.

Figure 5 shows a comparison of the electropherogram of AAV8 at the limit of detection ( $1.1 \times 10^9$  GC/mL) in 2 different formulation buffers. Red trace is AAV8 in 1X PBS which has a much higher salt content thus lower peak height, compared to AAV8 in 50 mM phosphate buffer. Additionally, the higher pH of phosphate buffer promotes the labeling reaction. In this figure the signal to noise



**Figure 4.** Comparison Between AAV8  $1.1 \times 10^{10}$  Using UV Absorption (red trace) and LIF (black trace). Left Y axis shows RFU units and Right Y axis show Absorbance. ? denotes an unknown peak.



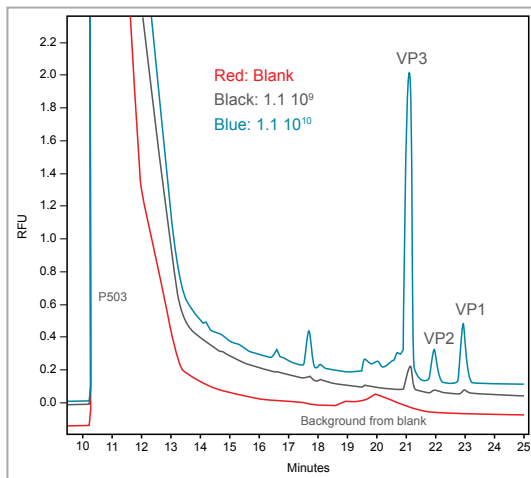
**Figure 5.** Comparison Between Different Salt Levels in Formulation Buffer of AAV8 at Low Trace Levels  $1.1 \times 10^9$  GC/mL. (Red) 1X PBS; (Black trace) 50 mM Phosphate buffer pH 8. \*marks an artifact from blank.

ratio were: 31, 3 and 7 for VP3, VP2 and VP1, respectively. This is within the acceptable S/N of 3 for limit of detection.

Additionally, the relative proportions between the capsids VP3, 2 and 1 were 10.8:1:1.8 respectively. While we demonstrated a separation of trace amounts of AAV with good signal to noise, this was accomplished with an increased amount of dye to

p3

[Download this technical note](#)



**Figure 6.** Overlay of Blank (red trace) and AAV8 at  $1.1 \times 10^9$  (black trace) and  $1.1 \times 10^{10}$  (blue trace). Note the raise in the blank run signifying a labeled product that cannot be attributed to an impurity in the AAV.

protein ratio than recommended by the manufacturer.<sup>3</sup> We found this necessary to promote labeling. However, we observed a significant tailing effect (Figure 5). This condition may be mitigated with buffer exchange and concentration of the initial AAV sample to  $10^{11}$  or higher GC/mL.

When working with labeled protein molecules, it is important to run a blank to ensure proper peak assignment. Even though P503 dye used in this work has very low quantum yield when unbound (<1%),<sup>3</sup> it is reactive towards primary amino groups and CE-SDS separation gel in a 100 mM Tris sample buffer. For this reason, a blank composed of AAV diluent buffer, formulation buffer and CE-SDS sample buffer was used. Figure 6 shows an overlay of the electropherograms from Blank and AAV8 at  $1.1 \times 10^9$  and  $10^{10}$  sample analysis.

Even though AAVs are stable, degradation is of great concern in any biopharma product. Figure 2 (see front page) showcases the comparison between heat stressed at  $93^\circ \text{C}$  for 30 minutes and non-stressed AAV8 samples at the same concentration  $1.1 \times 10^{10}$  GC/mL. Generally, there is a significant decrease in the overall peak intensity for the heat stressed samples. It is worth mentioning the considerable change in the profile of the electropherogram in the heat stressed samples specially for the lower molecular weight peaks presumed attributed to possible impurities in the original sample, are sensitive to the heat. Demonstrating that the proposed labeling scheme is suitable for assessing purity of AAV8 capsid proteins.

### Conclusions

- CE-SDS MW kit when combined with LIF detection can successfully achieve trace levels detection of AAV8 capsid proteins, successfully analyzing AAV8 at limit of detection at  $1.1 \times 10^9$  GC/mL.
- Easy 2 step labeling with P503 dye does not require buffer exchange prior to analysis.
- By using formulation buffer with low salt concentration and slightly higher pH, while still keeping AAV8 stability, allows for better sensitivity due to favorable electrokinetic injection conditions and improved sensitivity due to advantageous labeling environment.

### References

1. M. Agbandjie-McKenna et al.; *Molecular Therapy: Methods & Clinical Development*, (2017), 6, 171-182.
2. B. Bothner et al.; *Journal of Virology*, (2013), 87-24, 13150- 13160.
3. Chromeo P503 product insert, <https://www.activemotif.com/documents/1641.pdf>
4. Purity Analysis of Adeno-Associated Virus (AAV) Capsid Proteins using CE-LIF Technology; SCIEX Technical Note RUO-MKT-02-10086-A.
5. J. Fraser Wright, *Biomedicine* (2014), 2, 80-97.





# Nucleic acid sizing analysis

## Accurate Analysis of Double Stranded DNA over an Extended Size Range

### Capillary Electrophoresis-Laser Induced Fluorescence Detection (CE-LIF)

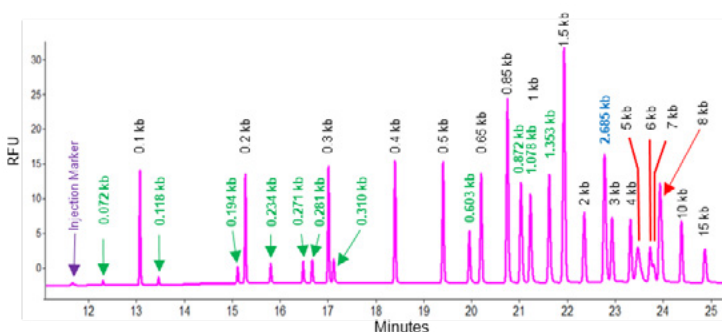
Jane Luo, Mukesh Malik, Marcia Santos, Sahana Mollah and Handy Yowanto  
SCIEX, Brea, California

#### Introduction

Analysis of double stranded DNA (ds DNA) is an integral part of manufacturing nucleic acids as therapeutics for gene and cell therapy as well as for DNA vaccines. The Food and Drug Administration (FDA) guidelines recommend manufacturers to include the gene map with relevant restriction sites for gene therapy vector constructs and vector diagrams identifying the gene insert, regulatory regions and pertinent restriction endonuclease sites when submitting applications for investigational new drugs (IND).<sup>1-2</sup> Analysis of restriction fragments from vectors of various sizes often requires accurate sizing over an extended size range. Traditional agarose gel method has limitation on long size coverage as well as poor fragment resolution, sizing accuracy and reproducibility. In this technical note, we describe CE-LIF based methods for accurate and robust analysis of ds DNA over an extended molecular size range from 100 base pairs (bp) to 15 kilo base pairs (kb).



**Figure 1.** The PA 800 Plus Pharmaceutical Analysis System (A), the dsDNA 1000 Kit (B) and the LIFluor Enhance Stain (C).



**Figure 2.** Example of successful separation and sizing of restriction fragments of virus and plasmid DNA using SCIEX dsDNA 1000 kit. Peaks with sizes labeled in green font are *Hae* III restriction fragments of  $\phi$ X 174 bacteriophage DNA. The linear pUC18dG plasmid DNA was labeled in blue font. Fragments in 1 kb plus DNA ladder are labeled in black font. The injection marker in purple font is LIF Performance Test Mix. Separation was performed on a 30 cm effective length DNA capillary at 7.8 kv voltage for 30 minutes.

#### Key Features

- Complete set of reagents for dsDNA fragment analysis with better kit shelf life than microchip based reagent kits
- Excellent resolution over an extended fragment size range with the same gel matrix
- Accurate size determination by standard curve
- Good Sensitivity (LOD of 0.79 ng/ml) and dynamic range (2.7 log)
- Good repeatability with % RSD of less than 0.5% for migration time and less than 2.5% for corrected peak area
- Ideal for restriction enzyme digestion analysis of DNA vectors

## Materials and Instrument

### Materials:

The eCAP dsDNA 1000 kit (PN 477410, Figure 1B), the LIFluor Enhance fluorescent stain (PN 477409, Figure 1C), dsDNA 1000 Test Mix (PN: 477414, Figure 1B), linearized pUC18dG (PN 608008) and LIF Performance Test Mix (PN: 726022) were from SCIEX, Framingham, MA. The 1 kb plus (PN 10787018) and 1 kb (PN 15615-016) DNA ladders were from Thermo, Carlsbad, CA. HPLC grade water (W5-4) was from Fisher Chemicals, New Lawn, NJ.

**Instrument and software:** A PA 800 Plus Pharmaceutical Analysis System (Figure 1A) equipped with LIF detector and solid-state laser with excitation wavelength at 488 nm and a 520 nm band pass emission filter were from SCIEX, Framingham, MA. Data acquisition and analysis were performed using 32 Karat software V10.

## Methods

**Gel buffer reconstitution:** This step must be done one day before running samples. To rehydrate the gel buffer, 20 ml of 0.2 µm filtered deionized water was added to the gel buffer vial. After the gel buffer vial was securely capped, the vial was placed on a rotator or the gel mixture was gently stirred with a small stirring bar for up to 24 hours or till the dried gel was completely dissolved. The hydrated gel should be good for one month if stored at 4°C.

**Cartridge Assembly:** DNA capillary (PN 477477) was installed per instructions on kit insert (PN 726412) in the dsDNA 1000 kit. The total capillary length was 40.2 cm with 30 cm as the length to the detection window. For analysis of large sized DNA fragments, a 50.2 cm total length capillary was also used with the length to the detection window as 40 cm. Since the inner wall of the DNA capillary is coated, the cartridge assembly was carried out within 5 to 10 minutes. The capillary ends were immersed in liquid (water or buffer) as soon as the cartridge assembly was complete to prevent the coating from drying out.

**Sample Preparation:** All samples were diluted with filtered deionized water to 2 to 20 ng/µl for analysis on the PA 800 Plus instrument.

**LIF Calibration:** To ensure consistent response of LIF detector throughout this study, the LIF detector was calibrated using LIF Calibration Wizard and Performance Test Mix (PN: 726022)

following the instructions in LIFluor Enhance user's guide (PN 725824). The target RFU value was set as 62. This calibration was done whenever capillary or LIF detector or the laser was changed. After the calibration was complete and the CCF was accepted, vials of water were placed back to the "home" positions (BI:A1 and BO:A1) so that the capillary ends were placed in water to keep them wet.

**Instrument Performance Test:** Instrument and gel buffer performance may be tested using dsDNA 1000 Test Mix (PN 477414) and LIFluor Enhance stain (PN 477409). Test Mix can be diluted by adding 1 mL of filtered deionized water (> 18MΩ) to the test mix vial. 5 µl LIFluor Enhance stain may be added to 6 mL of 1X rehydrated dsDNA 1000 gel buffer. Separation may be carried out following instructions in LIFluor Enhance user's guide (PN 725824).

**Preparation of Buffer Trays and Sample Trays:** Vial positions for buffer trays are indicated in Figure 3. Each "DDI Water" vials were filled with 1.5 ml double deionized (DDI) water. Waste vials were filled with 1 ml DDI water. "Gel Fill" vials and "Gel Sep." vials were filled with 1.5 ml gel. Vial "A6" on the outlet tray was filled with 1 ml gel. Two sets of water dip vials were used during sample separation for minimal sample carryover and better repeatability. Vial increment was done for every 8 injections. Injecting samples against a half-filled vial of gel buffer (BO:A6) facilitated improved precision in peak area.

Each ds DNA sample was transferred to a Microvial (PN 144709) at 50 to 100 µl. Sample vials were loaded onto the sample inlet tray.

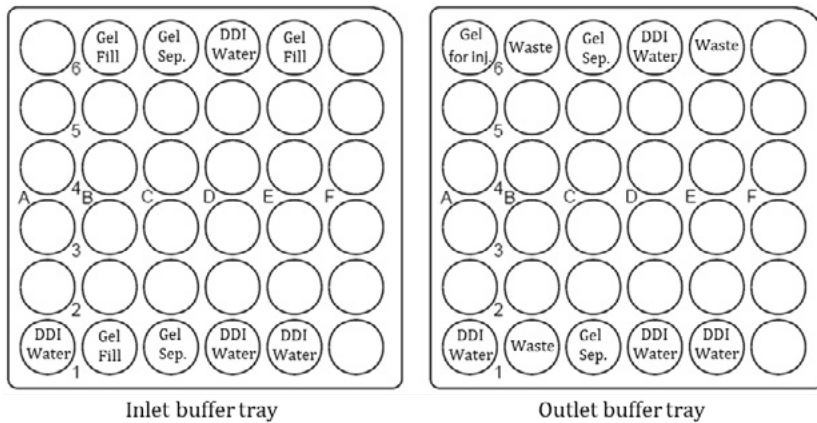
### Instrument Set up:

The "Initial Conditions" and "LIF Detector Initial Conditions" were set up as indicated in Figure 4 and Figure 5 respectively. Same set up was used for all three methods: Capillary Conditioning, Separation and Shut Down.

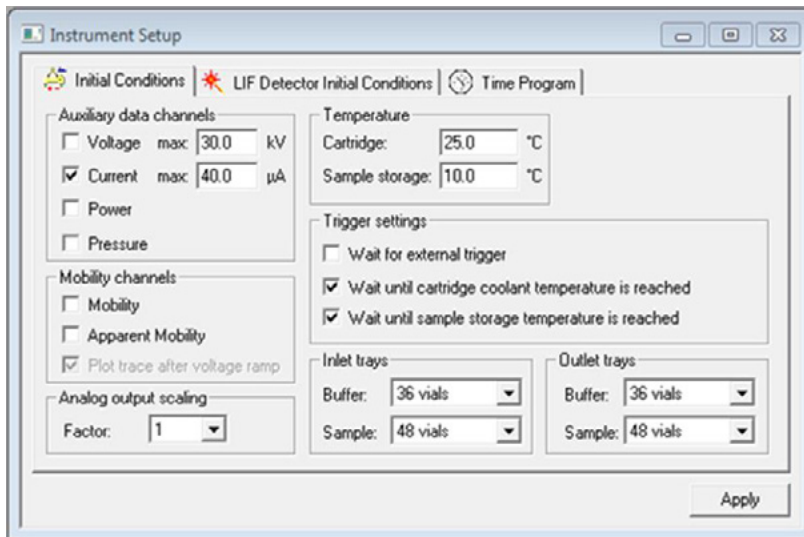
The time program for Capillary Conditioning was illustrated in Figure 6. Buffer vials used for conditioning were purposely positioned in row #6 to avoid moving the capillary tips over buffer vials used for sample separation.

Figures 7 and 8 show time programs for separation and shut down methods for running the dsDNA 1000 test mix. For DNA fragments with sizes larger than 2 kb, the separation time was extended to longer time as indicated in figure legends.

When the 40 cm effective length DNA capillary was used, injection condition was 0.5 psi for 10 sec. The separation time was 25 minutes at 12 kv, 35 minutes at 10 kv, 45 minutes at 8 kv, 65 minutes at 6 kv and 75 minutes at 5 kv.



**Figure 3.** Schematic set up for Inlet (left) and Outlet (right) buffer trays for up to 8 sample injections.



**Figure 4.** Settings for "Initial Conditions" tab.

[Download this technical note](#)

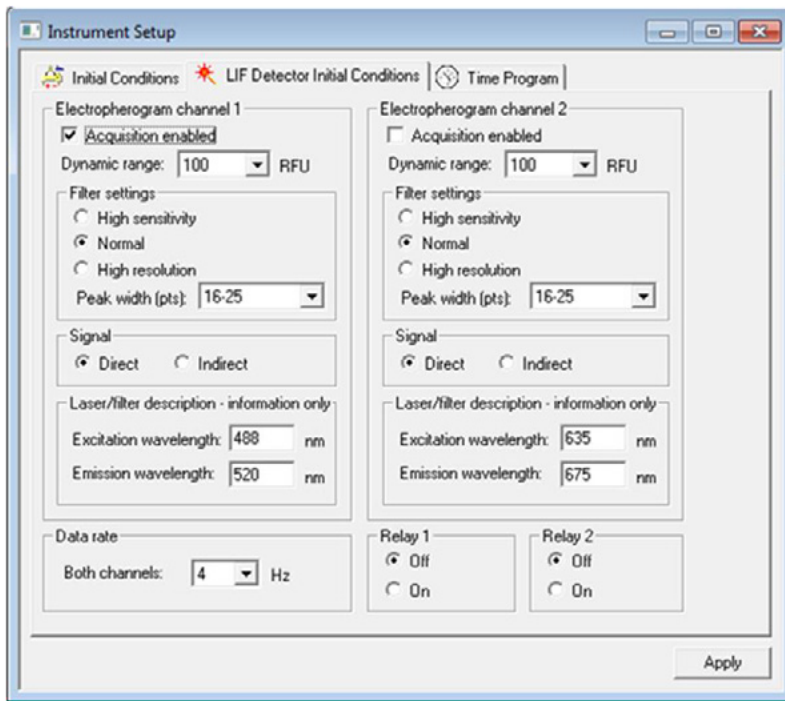


Figure 5. Settings for “LIF Detector Initial Conditions” tab.

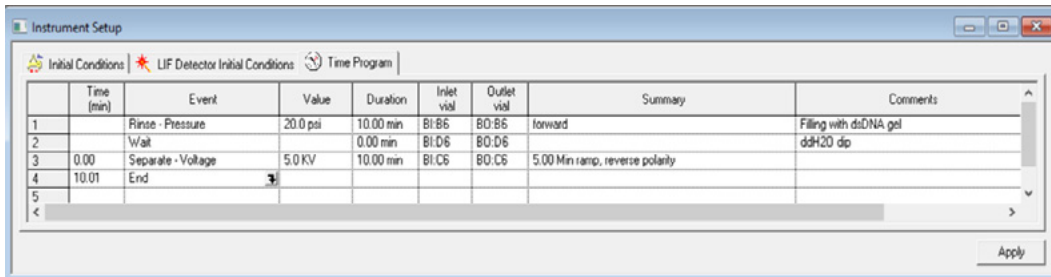
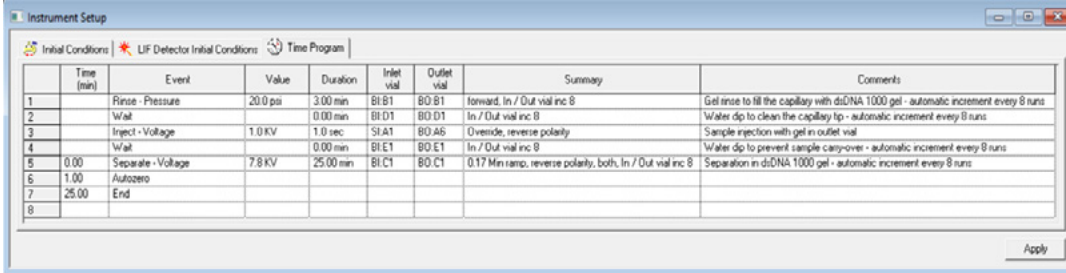


Figure 6. Time Program Settings for the Conditioning Method.

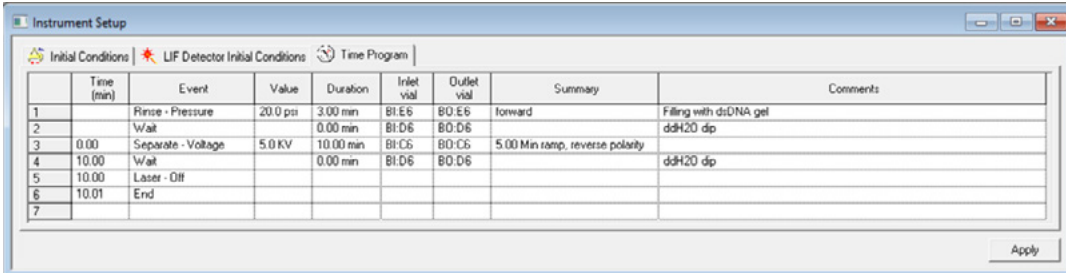
[Download this technical note](#)





Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1	Rinse - Pressure	20.0 psi	3.00 min	BI:B1	BO:B1	forward, In / Out vial inc 8	Get rinse to fill the capillary with dsDNA 1000 gel - automatic increment every 8 runs
2	Wait		0.00 min	BI:D1	BO:D1	In / Out vial inc 8	Water dip to clean the capillary tip - automatic increment every 8 runs
3	Inject - Voltage	1.0 KV	1.0 sec	SI:A1	BO:AS	Overside, reverse polarity	Sample injection with gel in outlet vial
4	Wait		0.00 min	BI:E1	BO:E1	In / Out vial inc 8	Water dip to prevent sample carryover - automatic increment every 8 runs
5	Separate - Voltage	7.8 KV	25.00 min	BI:C1	BO:C1	0.17 Min ramp, reverse polarity, both, In / Out vial inc 8	Separation in dsDNA 1000 gel - automatic increment every 8 runs
6	1.00	Autzero					
7	25.00	End					
8							

Figure 7. Time Program Settings for the Separation Method of the dsDNA 1000 Test Mix.



Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments	
1	Rinse - Pressure	20.0 psi	3.00 min	BI:E6	BO:E6	forward	Filling with dsDNA gel	
2	Wait		0.00 min	BI:D6	BO:D6		ddH2O dip	
3	0.00	Separate - Voltage	5.0 KV	10.00 min	BI:C6	BO:C6	5.00 Min ramp, reverse polarity	
4	10.00	Wait		BI:D6	BO:D6		ddH2O dip	
5	10.00	Laser - Off						
6	10.01	End						
7								

Figure 8. Time Program Settings for the Shut Down Method.

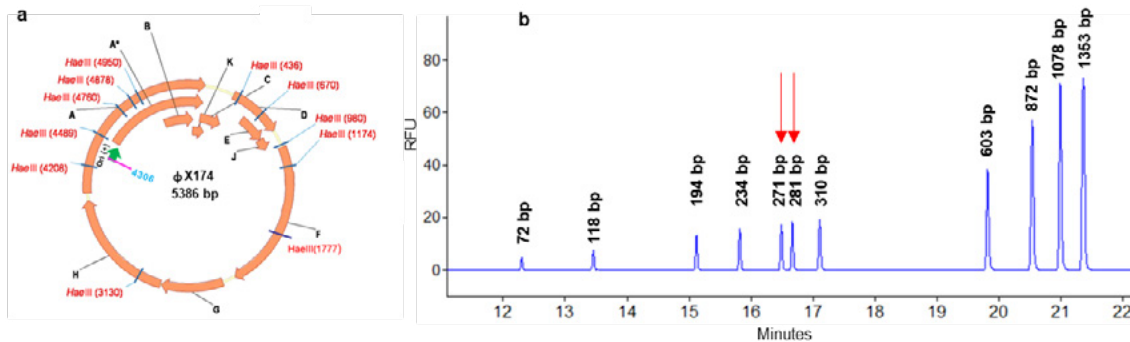
## Results and Discussion

### Analysis of the *Hae* III restriction fragments of $\phi$ X174

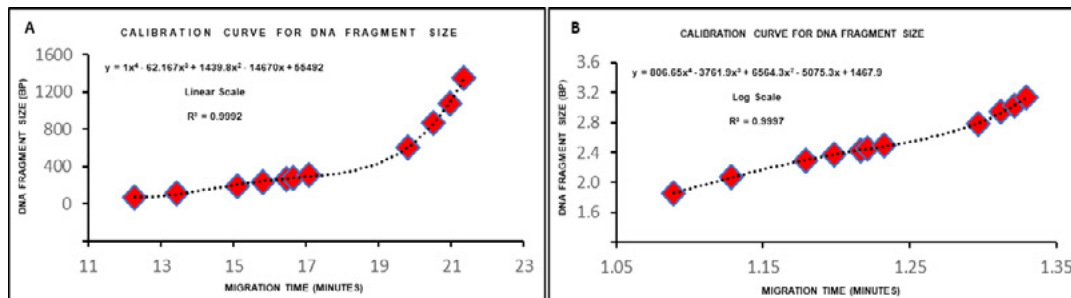
The  $\phi$ X174 is a small, single-stranded DNA virus that infects *Escherichia coli*. It is also called a bacteriophage. It was the first genome to be sequenced.<sup>3</sup> Its replicative form (RF) is a double stranded circular DNA molecule of 5386 base pairs (bp). When it is digested by a restriction enzyme *Hae* III, 11 DNA fragments are generated.<sup>4</sup> These fragments have been used as popular DNA standards for size determination of DNA fragments in the past 40 years. It is available from SCIEX as dsDNA 1000 Test Mix (PN 477414). Figure 9a shows the restriction map of  $\phi$ X174 with *Hae* III sites, the origin of replication, and location of 11 genes indicated. Figure 9b is a representative electropherogram obtained by separation of the *Hae* III fragments of  $\phi$ X174 on

PA800 Plus using dsDNA 1000 gel, LIFluor Enhance dye and coated DNA capillary at 40.2 cm total length. Instrument set up and separation conditions were described in the Methods section. All 11 fragments including two fragments differing by 10 bps were baseline resolved within a 25 minutes run. As shown in Figure 10, a calibration curve was generated with results from Figure 9b. The best curve fitting was obtained with the quartic (4<sup>th</sup> degree polynomial) model with excellent R square values: 0.9992 at the linear scale (panel A) and 0.9997 at log scale (panel B).

Further experiments were carried out to determine the limit of detection (LOD) and limit of quantitation (LOQ) as well as the linearity of detector response to concentration of the DNA sample in this method. The *Hae* III fragments of  $\phi$ X174 were serially diluted by 2 fold each time and separated. Figure 11A shows the electropherograms obtained with some of the diluted samples. The concentration of the 872 bp fragment is indicated next to each electropherogram trace. The signal to noise ratio for the 872 bp peak was 5.4 at the concentration of 0.79 ng/ml, and



**Figure 9.** Baseline resolution of 11 *Hae* III restriction fragments of  $\phi$ X174. Panel "a" shows the map of  $\phi$ X174 genome with *Hae* III sites indicated. The  $\phi$ X174 genome contains 11 genes (A, A\*, B-H, J, K) represented by thick arrows in orange color. Location of the origin of replication is indicated by the green arrow. Panel "b" shows an electropherogram of the 11 *Hae* III fragments with their sizes labeled in base pairs (bp). Separation was performed with a 30 cm effective length DNA capillary (40.2 cm in total length), dsDNA 1000 gel and the LIFluor Enhance stain. The two red arrows indicate that the two fragments with 10 bp size difference were baseline resolved.



**Figure 10.** Calibration curve for DNA fragment size determination using dsDNA 1000 test mix ( $\phi$ X174 *Hae* III restriction fragments) as standards. DNA fragment sizes in bp were plotted against migration time in minutes. The best curve fitting model was quartic (4<sup>th</sup> degree polynomial). The R square value was 0.9992 for the linear scale (Panel A) and 0.9997 for the log scale (Panel B).

16.1 at 1.6 ng/ml. Therefore, the LOD is 0.79 ng/ml and the LOQ is around 1 ng/ml. Figure 11B shows that good linearity was obtained when plotting the corrected peak area of the 872 bp fragment against its concentration from 0.79 ng/ml to 404 ng/ml. The R square value was 0.995. The dynamic range was 2.7 log.

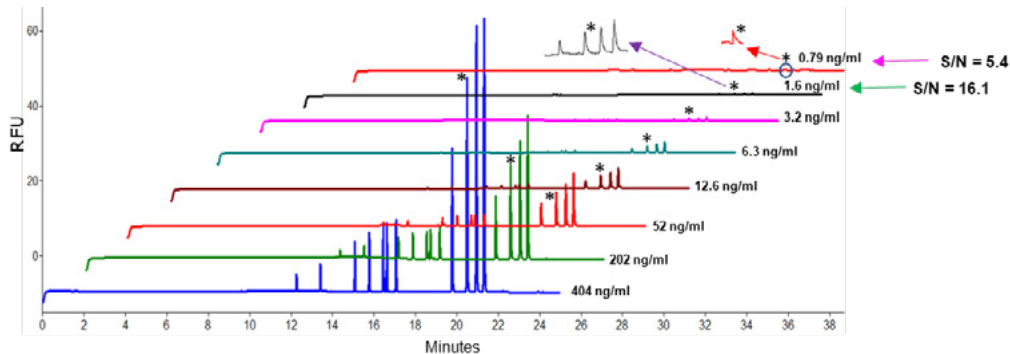
#### Analysis of DNA fragments over extended size range

Many of the vectors used in producing nucleic acid therapeutics for gene and cell therapy as well as for DNA vaccination have sizes well over 10 kb. Restriction enzyme digestion of these

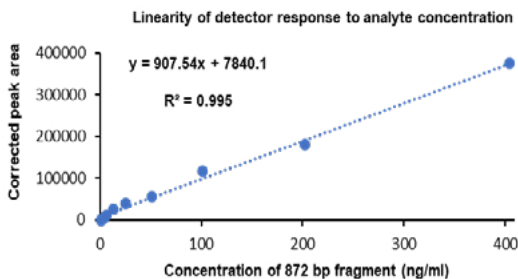
large vectors often generates some fragments that are larger than 1 kb as well as fragments that are smaller than 500 bps. Analysis of these fragments over an extended size range requires the use of DNA size standards that provide a long size range coverage. Several DNA standards with broader range of size coverage were tested. Among them, the 1 kb DNA ladder and 1 kb plus DNA ladder from Thermo were selected.

#### Analysis with 1 kb DNA ladder

The 1 kb DNA ladder includes 12 large fragments created by repeating a 1018 bp fragment for 1 to 12 times. It also contains additional 11 fragments with sizes from 75 bp to 1636 bp that are



**Figure 11A.** Good detector response to different concentrations of the sample. The dsDNA 1000 Test Mix was serially diluted before analysis. The 872 bp fragment used for quantitative analysis was labeled with “\*”. Its concentration in each sample is labeled.



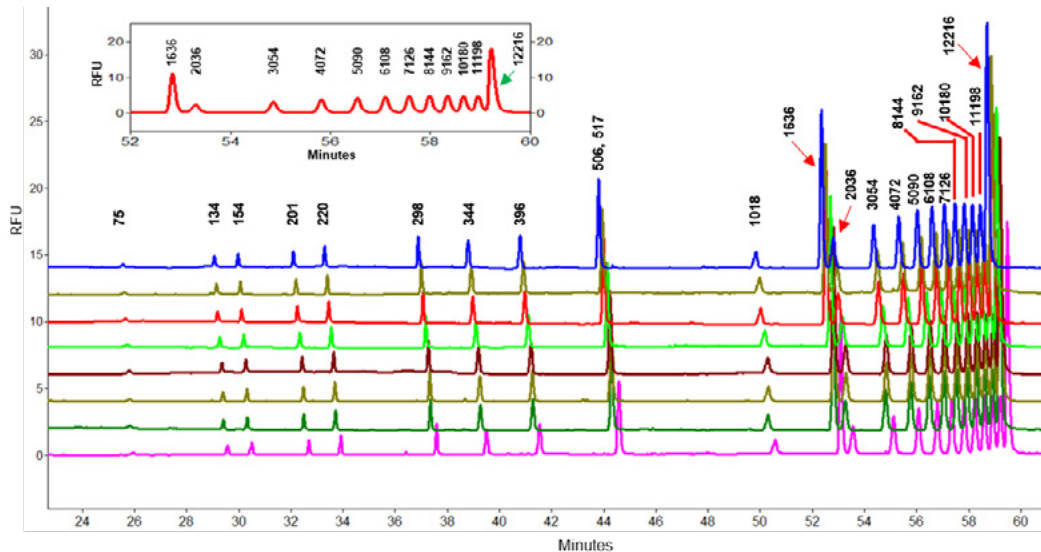
**Figure 11B.** Good linearity between detector response to analyte concentration. Corrected peak area of the 872 bp fragment was plotted against its concentration in ng/ml.

generated by *Hinf* I digestion of a plasmid vector pBR322. For the 1 kb DNA ladder, the best separation was obtained with the separation voltage at 6 kv and a 40 cm effective length DNA capillary. As shown in Figure 12, an experiment was done to test method repeatability. The 1 kb DNA ladder at the concentration of 20 ng/ul was injected at 0.5 psi for 10 sec for 8 runs and separated using dsDNA 1000 gel and LIFluor Enhance dye. Each separation was performed at a voltage of 6 kv for 65 minutes with 20 psi pressure at both capillary ends. Results in Figure 12 showed that the peak patterns were consistent between the 8 different runs. Fragments under 500 bp with size differences of 20 to 50 bp were well separated. The picture inset

in Figure 12 shows a zoomed-in view of the area where large DNA fragments with sizes of 1.6 kb to 12 kb were baseline separated as well. Good repeatability was demonstrated by the RSD% values for the 1636 bp peak: 0.43% for migration time and 2.19% for corrected peak area.

#### Analysis with 1 kb plus DNA ladder

The 1 kb plus DNA ladder contains 18 fragments in the range of 100 bp to 15,000 bp. Two different methods were evaluated with this DNA ladder. The first one included using the DNA capillary at 30 cm effective length and electrokinetic sample injection. An experiment was done to simulate analysis of fragments produced by restriction enzyme digestion of a plasmid and a virus. The 1 kb plus DNA ladder (at 6.25 ng/ul), linearized plasmid pUC18dG (at 3 ng/ul) and *Hae* III fragments of  $\phi$ X174 (at 2.5 ng/ul) were injected at 1 kv for 2 seconds. Separation was performed at 7.8 kv for 30 minutes. Results in Figure 2 demonstrated that excellent resolution was achieved with baseline resolution of fragments differing by 6-10 bp around 200-300 bp region, 30-50 bp around 800 bp region and 150 to 315 bp around 1 to 3 kb area. A calibration curve was created by plotting the size of DNA fragments in the 1 kb plus ladder against migration time using the quantitative analysis feature of the 32 Karat software. The R square value was 0.9982 at the log scale using the quartic model (data not shown). This calibration curve was used to deduce predicted sizes for the linearized pUC18dG and 10 of the *Hae* III



**Figure 12.** Good repeatability in separation of DNA fragments in the 1 kb DNA ladder. The 1 kb DNA ladder at the concentration of 20 ng/ul was injected 8 times and separated on PA800 Plus using dsDNA 1000 gel and LIFluor Enhance dye. Each separation was performed at a voltage of 6 kv for 65 minutes with 20 psi pressure at both capillary ends. The RSD% for the 1636 bp fragment was 0.43% for migration time and 2.19% for corrected peak area. The picture insert shows a zoomed-in view of the area where large DNA fragments with sizes of 1.6 kb to 12 kb were baseline resolved. The effective length of the DNA capillary used was 40 cm.

fragments of  $\phi$ X174. The predicted sizes were compared to their corresponding theoretical sizes (Table 1). The differences between predicted sizes and theoretical sizes were no more than 5 to 7% of the theoretical sizes for all fragments, demonstrating the accuracy of size determination for restriction fragments by this method. In addition, this method can be potentially useful for analysis of host cell DNA in cell culture-produced vaccines in which over 80% of residual DNA was under 1000 bp with the major peak around 150 bp and some minor peaks in the range of 1000 bp to 10 kb.<sup>5-6</sup>

The second method with the 1 kb plus DNA ladder involves using the DNA capillary at 40 cm effective length and pressure injection for sample loading. Figure 13 shows results obtained by injecting a mixture of the 1 kb plus DNA ladder and the 1 kb DNA ladder at 0.5 psi for 10 seconds. Separation was carried out at 10 kv for 35 minutes. Large DNA fragments were separated with a resolution of 1 kb. In the region between 850 to 2000 bp, the resolution was 136 to 150 bp. Fragments smaller than 500 bp were well resolved when the size differences between them were 20 bp or larger. Interestingly, a 298 bp fragment was separated

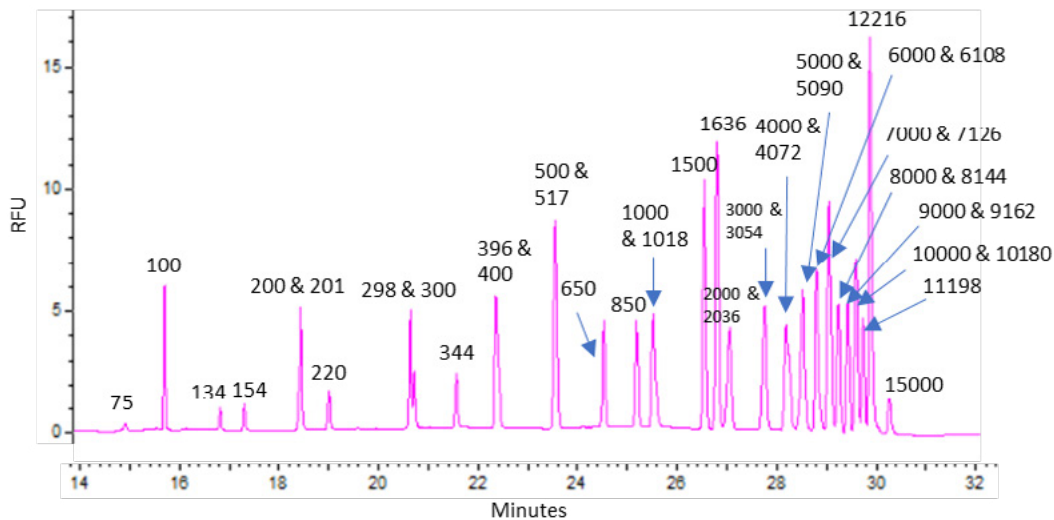
**Table 1.** Comparison of theoretical sizes with predicted sizes for 11 fragments.

Theoretical Size	118	194	234	271	281	310	603	872	1078	1353	2685
Predicted Size	120	204	236	265	270	295	609	931	1054	1331	2804
$\Delta$ Size	2	10	2	-6	-11	-15	6	59	-24	-22	119
% $\Delta$ Size / TS	2	5	1	-2	-4	-5	1	7	-2	-2	4

Notes: " $\Delta$  Size" is the difference between theoretical size and the size predicted using 1 kb plus DNA ladder. "%  $\Delta$  Size / TS" stands for percent size difference relative to the theoretical size.

from a 300 bp fragment while a 396 bp fragment was not resolved from a 400 bp fragment. These differences may be due to differences in sequence composition or differences in interaction with the intercalating dye. Resolution achieved with this method was similar to that obtained with the first method.

Users can choose which method to use based on their samples.



**Figure 13.** Good resolution was achieved in separating fragments from a mixture of the 1 kb DNA ladder and the 1 kb plus DNA ladder. Samples containing 20 ng/ $\mu$ l of both DNA ladders were injected at 0.5 psi for 10 seconds. Separation was performed with a 40 cm effective length DNA capillary at 10 kv for 35 minutes.

## Conclusions

This technical note demonstrated the following advantages of using SCIEX dsDNA 1000 gel, coated DNA capillary and LIFluor Enhance dye over the traditional agarose gel and microchip based methods in analysis of dsDNA fragments:

1. Accurate size determination is achieved over a longer size range (72 bp to 15 kb) with the same dsDNA 1000 gel matrix and within 35 minutes
2. Finer resolution for more accurate size determination
  - a. 6 to 10 bp around 200-300 bp
  - b. 30 to 50 bp around 800 bp to 1kb
  - c. 136 to 315 bp around 1.5 kb to 3 kb
  - d. 0.5 kb to 1 kb around 4 kb to 12 kb
3. Better sensitivity and assay repeatability for more consistent results with LOD of 0.79 ng/ml and %RSD of less than 0.5% for migration time and less than 2.5% for corrected peak area
4. Wider dynamic range (2.7 log) allows detection of both abundant fragments and rare fragments
5. Longer kit shelf life (15 months vs 4 to 6 months) provides cost-effectiveness and efficient inventory management

In addition, co-injection of sample and DNA standards provides higher accuracy in size determination than traditional agarose gel electrophoresis.

Furthermore, users can choose different capillary length, DNA standards, sample injection and separation conditions to suit their analysis needs of different size range.



### References

1. Guidance for FDA reviewers and sponsors. Content and review of chemistry, manufacturing, and control (CMC) information for human gene therapy investigational new drug applications (INDs). 2008.
2. Content and review of chemistry, manufacturing, and control (CMC) information for human gene therapy investigational new drug applications (INDs). FDA Draft Guidance to Industry. 2018.
3. Sanger, F. *et al.* (1978) The nucleotide sequence of bacteriophage phiX174. *J. Mol. Biol.* **125**:225-246.
4. Middleton, *et al.* (1972) Specific fragments of phi X174 deoxyribonucleic acid produced by a restriction enzyme from *Haemophilus aegyptius*. endonuclease Z. *J. Virology.* **10**:42-50.
5. FDA. Guidance to industry. Characterization and qualification of cell substrates and other biological materials used in the production of viral vaccines for infectious disease indications. 2010.
6. Shen *et al.* (2013) Size analysis of residual host cell DNA in cell culture-produced vaccines by capillary gel electrophoresis. *Biologicals* **41**:201-208.

AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-9967-A 07/2019



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



## Method Evaluation for RNA Purity Analysis Using CE-LIF Technology

Tingting Li, Mukesh Malik, Handy Yowanto  
SCIEX Separations, Brea, CA

### Introduction

CE has been demonstrated to be one of the most powerful techniques for analysis of a wide variety of molecules. The PA800 Plus Pharmaceutical Analysis System from SCIEX is an outstanding instrument, specifically designed and optimized for capillary gel electrophoresis (CGE) separation, which is widely used for separation of biologics such as protein, peptide, nucleic acids, etc.

Comparing to traditional slab-gel based electrophoresis methods, CGE offers superior resolution, shorter analysis time, automated operation and exceptional sensitivity when combined with a laser induced fluorescence detector. Compared to chip-based CE systems, the PA800 Plus provides open chemistry, which enables flexibility for method modification and optimization to generate optimal results for each specific project.

Recent advances in gene therapy research have gained promise in the utility of gene therapeutics compounds. The ability to quantify RNA purity and quality is critical to ensure the safety and efficacy of these molecules. In this study, a CE-LIF fast separation method which was used to evaluate total RNA quality<sup>1</sup> was optimized to achieve higher resolution; and the method was then evaluated for RNA purity determination. The resolution of this method was optimized for RNA size ranging from 200 bases to 6583 bases and evaluated by spiking a 1.2 kb positive RNA marker into several commercially available RNA ladders. A calibration curve was generated from the RNA ladders which could then be used to estimate the size of unknown sample peaks. In addition, assay repeatability, linearity, LOQ, LOD were also evaluated in this study. The optimized method could be used as an RNA platform analytical method for RNA analysis or further modified to suit more specific user criteria.

### Key Features

- Baseline resolution of RNA ladders from 0.2 kb to 6.5 kb
- Capability of RNA size estimation of unknown samples using calibration curve of RNA ladders



The PA 800 Plus Pharmaceutical Analysis System

- Good repeatability with %RSD of MT within 0.1% and %RSD of %CPA within 5%
- Excellent Linearity with R<sup>2</sup>=0.996 of detection response vs concentration
- LOQ = 0.33 ng/mL and LOD = 0.081 ng/mL

### Materials and Methods

Urea was obtained from ThermoFisher Scientific (PN 29700, Waltham, MA). Polyvinylpyrrolidone was obtained from Sigma-Aldrich (PN 437190, St Louis, MO). TBE Buffer, 10X, Molecular Biology Grade was obtained from Sigma-Aldrich (PN 574795, St Louis, MO). SYBR™ Green II RNA Gel Stain, 10,000X concentrate in DMSO was obtained from ThermoFisher Scientific (PN S-Waltham, MA), Ladder of Nine RNA Transcripts 281-6583 bases was obtained from Promega (PN G3191, Fitchburg, WI), 1.2 kb Kanamycin Positive Control RNA was obtained from Promega (PN C1381, Fitchburg, WI).

1%PVP (1.3 MDa) in 1X TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) with 4 M urea and 50,000x dilution or 0.002% SYBR green dye was used as separation buffer.

RNA ladders and 1.2 k RNA marker were diluted in DDI water, denatured for 5 minutes at 65° C and cooled down on ice

p1

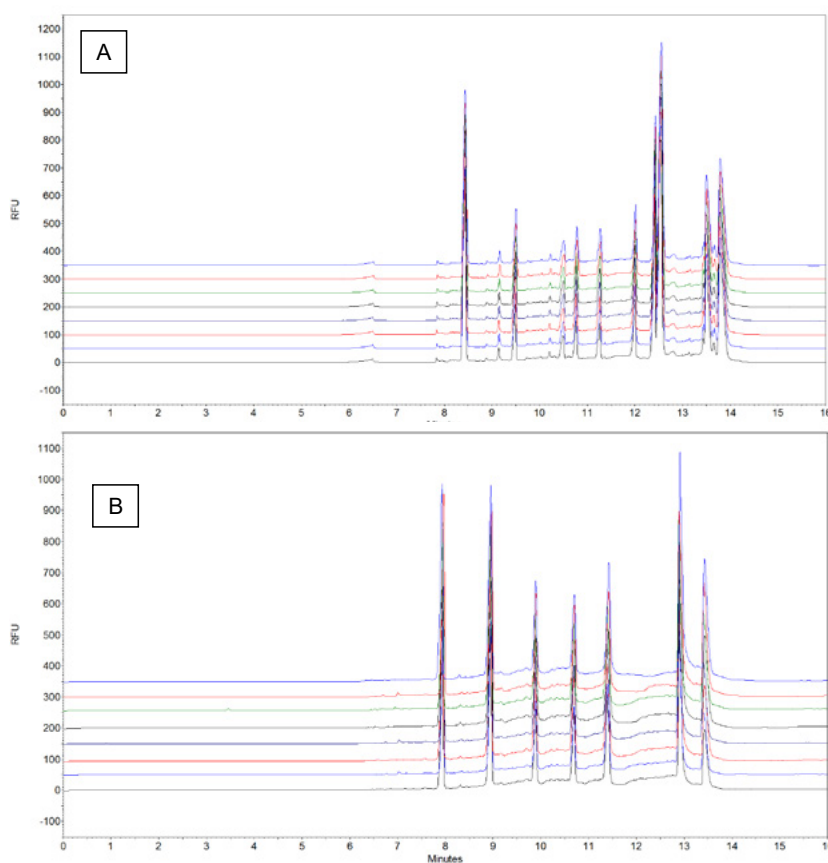
[Download this technical note](#)

for 5 minutes before loading. Two RNA ladders were used in this study. One is the Thermo RNA 6000 Ladder of six RNA transcripts with lengths of 0.2, 0.5, 1.0, 2.0, 4.0, and 6.0 kb. The other one is the Promega RNA marker which consist of a ladder of nine RNA transcripts of 281, 623, 955, 1,383, 1,908, 2,604, 3,638, 4,981 and 6,583 bases. Both ladders were diluted to 25 µg/mL and spiked with the proper amount of the 1.2 kb marker, respectively.

All experiments were performed on PA800 plus Pharmaceutical Analysis System from SCIEX.

EZ cartridge pre-assembled with bare fused-silica capillary (50 µm I.D., 30 cm total length, 20 cm effective length) was purchased from SCIEX (PN A55625, Framingham, MA).

Samples were introduced into the inlet of the capillary either electrokinetically at -5kV for 3s. Separations were performed using reversed polarity with 300V/cm electrical field at 25° C. Samples Tray were kept at 4° C to minimize RNA degradation and renaturation. LIF detector was configured with a 488-nm laser with an emission filter of 520 nm.



**Figure 1.** a) The overlay of the eight consecutive analyses of the Promega RNA ladder with spiked in 1.2 kb marker. b) The overlay of the eight consecutive analyses of the Thermo RNA 6000 ladder with spiked in 1.2 kb marker.

p2

[Download this technical note](#)

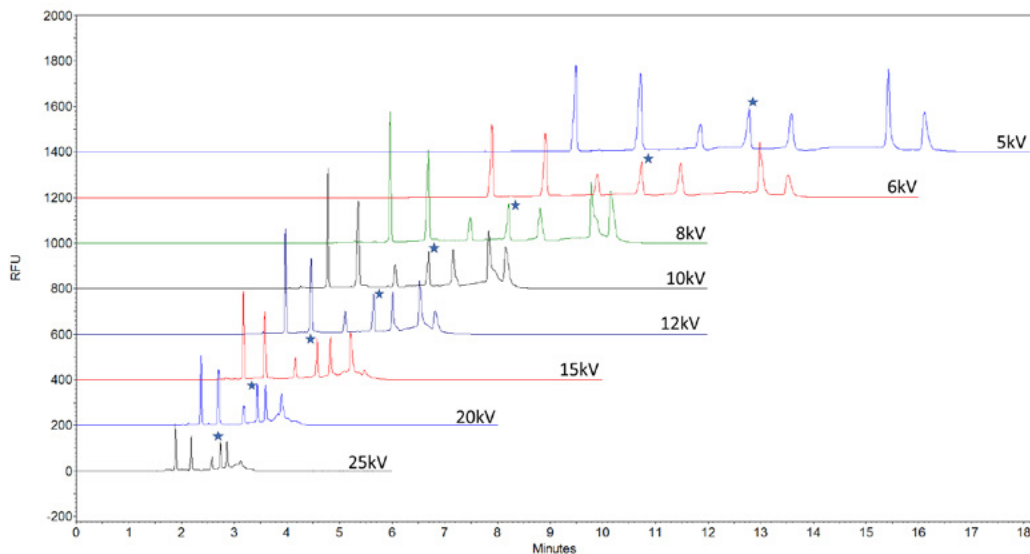
## Results and Discussion

### Method optimization

Both the Thermo RNA6000 ladder with spiked in 1.2 kb marker and Promega ladder with spiked in 1.2 kb marker were used for method optimization as the Thermo RNA 6000 ladder has less impurity peaks and Promega ladder has more markers of different sizes in the mixture. Different separation voltages and separation temperature were compared to find the optimal resolution and peak shape. Figure 2 compares the separation of Thermo RNA6000 ladder (spiked in 1.2 kb marker) using 9 different separation voltages (25 kV, 20 kV, 15 kV, 12 kV, 10 kV, 8 kV, 6 kV and 5 kV) and a capillary temperature of 25° C. All The 6 markers in the RNA 6000 ladder and a spiked in 1.2kb RNA molecule were baseline resolved using 6 kV and 5kV of separation voltage at 25° C. Figure 3 shows the similar study using RNA ladder from Promega, which has 9 markers, including a spiked in 1.2 kb molecule. All 9 RNA markers (281 bases, 623 bases, 955 bases, 1.2 kb,1383 bases, 1908 bases, 2604 bases, 3638 bases, 4981 bases and 6583 bases) were baseline resolved using 8 kV and 6 kV of separation voltage

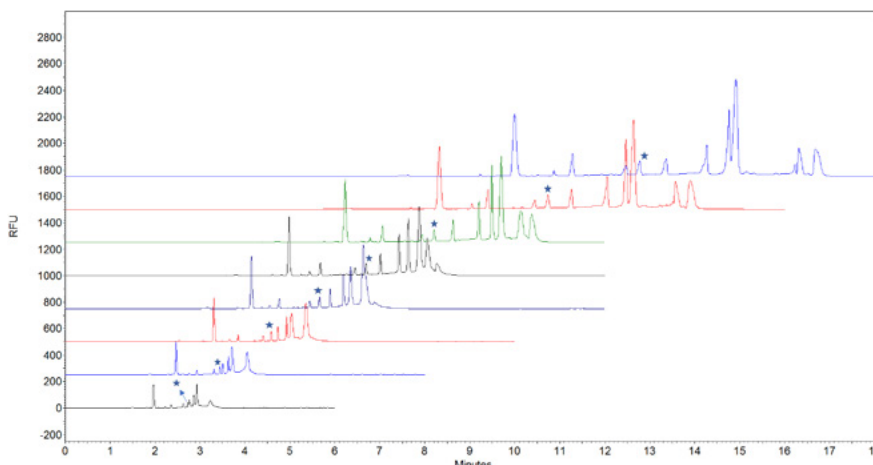
at 25° C. A separation voltage of 6 kV and capillary temperature of 25° C provided the best resolution and peak shape for these two different RNA markers.

Similar optimization work was performed using capillary cartridge temperatures of 15° C, 20° C, 30° C, 35° C and 40° C (data not shown). Of all the cartridge temperature and electric field conditions tested, 6 kV (200 v/cm) at 25° C was considered the optimized condition when comparing factors such as analysis time, resolution, peak shape, and stability of RNA sample solutions. Additional work was done to evaluate other parameters such as 2% PVP in the separation buffer, injection method (pressure injection for 3s at 0.5 psi, electrokinetic injection for 5s at 2 kV and 3s at 5 kV) and capillary length (10 cm and 20 cm effective length options provided by EZ cartridge). The following optimized method conditions were determined: separation buffer containing 1% PVP, electrokinetic injection for 3s at 5 kV, 20 cm effective length capillary and separation with 200 V/cm electrical field (6 kV) at 25° C. These method parameters were used in the evaluation study.



**Figure 2.** Overlay traces of RNA 6000 ladder with spiked in 1.2 kb marker separated using different electric field under 25° C. The 6 RNA markers in each trace are 0.2 kb, 0.5 kb, 1.0 kb, 2.0 kb, 4.0 kb and 6.0 kb. The peak with star mark is the spiked in 1.2 kb molecule. The separation voltages used from bottom to top traces were 25 kV, 20 kV, 15 kV, 12 kV, 10 kV, 8 kV, 6 kV and 5 kV.



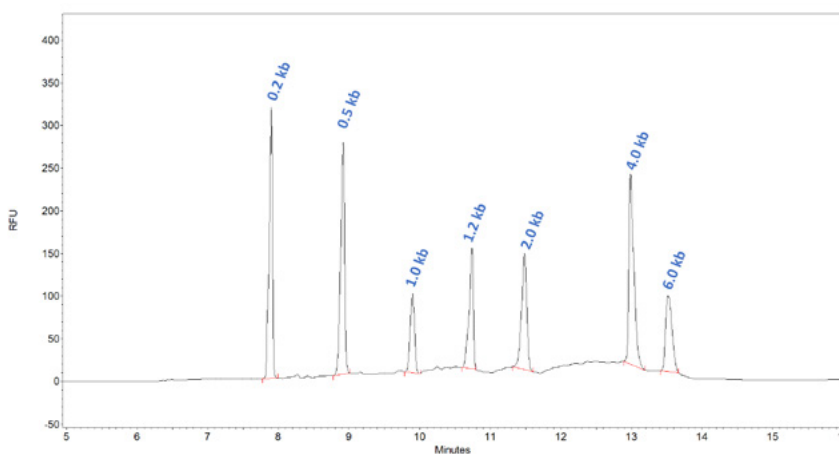


**Figure 3.** Overlay traces of Promega RNA ladder with spiked in 1.2kb marker separated using different electric field under 25° C. The 9 RNA markers in each trace are 281 bases, 623 bases, 955 bases, 1383 bases, 1908 bases, 2604 bases, 3638 bases, 4981 bases and 6583 bases. The peak with star mark is the spiked in 1.2 kb molecule. The separation voltages used from bottom to top traces were -25 kV, -20 kV, -15 kV, -12 kV, 10 kV, 8 kV, 6 kV and 5 kV.

**Calibration Curve**

Figures 4 and 5 illustrate the typical electropherograms of the analysis of the RNA6000 ladder and Promega RNA Ladder using the optimized method conditions. Figures 6 and 7 illustrate plots of the log RNA size (in bases) versus Migration and demonstrate good linearity of response of each ladder. Base numbers were used to plot the calibration curve instead of MW

since the MW information was not provided from the vendor of the ladders. The coefficient of determination ( $R^2$ ) of RNA6000 ladder is 0.993273 while the  $R^2$  of Promega RNA ladder is 0.988811, which is slightly less than 0.99. One of the possible reasons for the relatively low  $R^2$  value of the calibration curves is the different composition of the nucleic bases in the sequence of each RNA marker, which leads to the deviations of the relationship of MW and base numbers of the RNA markers.



**Figure 4.** E-gram of the Thermo RNA 6000 Ladder with spiked in 1.2kb marker separated by the optimized method.

p4

[Download this technical note](#)

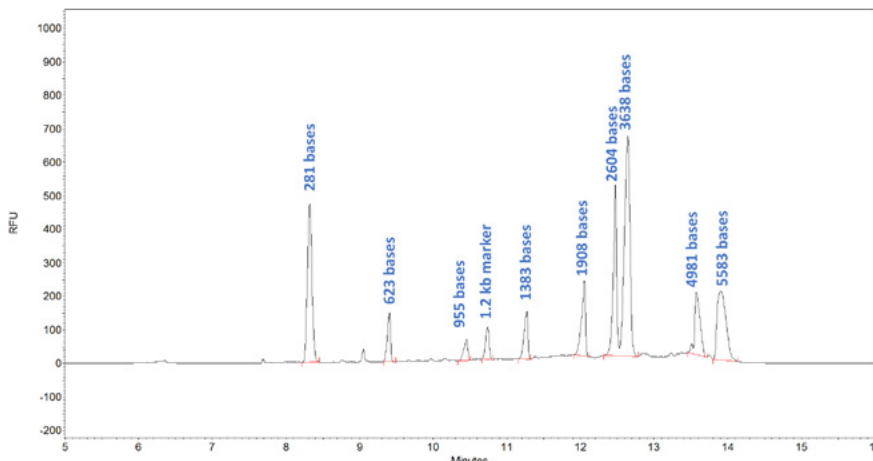


Figure 5. E-gram of the Promega RNA ladder spiked in 1.2 kb marker separated by the optimized method.

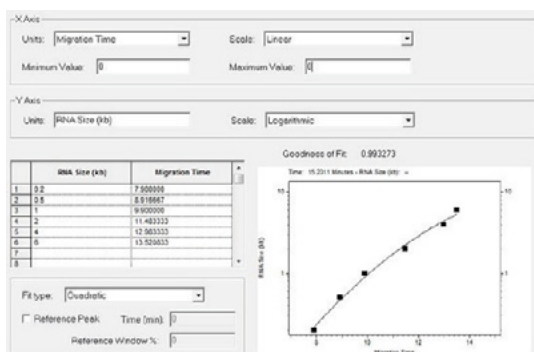


Figure 6. Calibration curve of log RNA Size in kb vs migration time of the Thermo RNA 6000 Ladder. R2 is 0.993273.

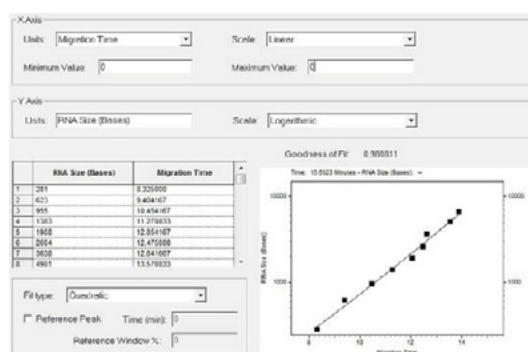


Figure 7. Calibration curve of log RNA Size in bases vs migration time of the Promega RNA ladder. R2 is 0.988811.

**Repeatability**

Both RNA ladder samples were injected eight times to demonstrate the repeatability of the method. Figure 1a and 1b show the overlay of the eight consecutive analyses of the Promega RNA ladder with spiked in 1.2 kb marker and the Thermo RNA6000 Ladder with spiked in 1.2 kb marker, respectively. Table 1 summarizes the results of eight consecutive analyses of the Promega RNA ladder spiked in 1.2 kb marker. The relative standard deviation (%RSD) of the mobility for each

RNA markers in the Promega RNA sample was < 0.1%, while the %RSD of the quantitative determination of the % corrected area of each marker was < 1.5%. Table 2 summarizes the results of repetitive injections of the Thermo RNA6000 ladder with spiked 1.2 kb marker. The relative standard deviation (%RSD) of the mobility of each RNA marker in the Thermo RNA6000 ladder sample was < 0.15%, while the %RSD of the % corrected area of each marker in the Thermo RNA6000 ladder sample was < 5%. The results demonstrate the excellent repeatability of this method.

	RSD%	
	CPA%	MT
281 bases	0.82	0.08
623 bases	1.26	0.09
955 bases	1.22	0.07
1200 bases	1.15	0.09
1383 bases	1.25	0.06
1908 bases	1.37	0.07
2604 bases	0.95	0.09
3638 bases	0.59	0.06
4981 bases	1.16	0.04
6583 bases	0.91	0.04

**Table 1.** The %RSD of the Migration Time (MT) and Corrected Peak Area Percentage (CPA%) of the 9 markers and a spiked in 1.2kb RNA molecule.

	RSD%	
	CPA%	MT
0.2 kb	2.13	0.15
0.5 kb	1.10	0.09
1.0 kb	1.46	0.07
1.2 kb	2.91	0.05
2.0 kb	3.01	0.11
4.0 kb	2.58	0.09
6.0 kb	4.68	0.09

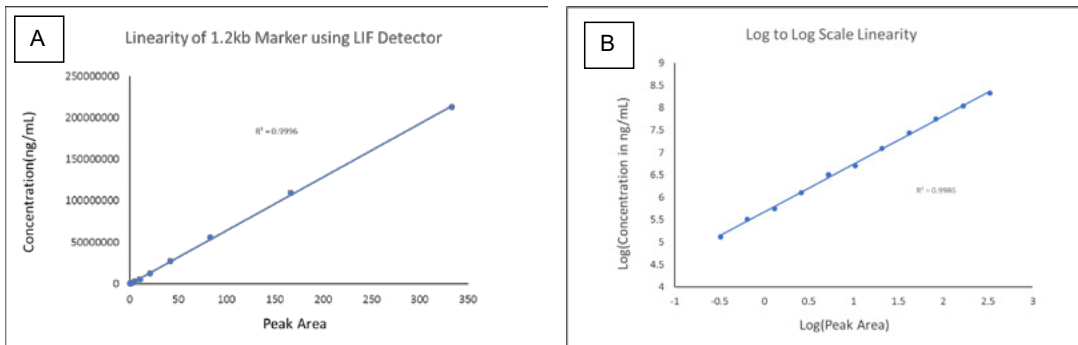
**Table 2.** The %RSD of the Migration Time (MT) and Corrected Peak Area Percentage (CPA%) of the 6 markers and a spiked in 1.2kb RNA molecule.

**Linearity, LOD and LOQ using LIF Detector**

The 1.2 kb Positive Control RNA molecule with a known concentration of 0.5mg/mL was used for linearity, LOD and LOQ studies.

Figure 8 demonstrates the excellent linearity of detection

response using the LIF detector, plotting peak area versus concentration (ng/mL) over the range of 0.33 ng/mL to 333 ng/mL for the 1.2 kb RNA Marker. The R2 is 0.9996. The LOQ of 1.2 kb RNA marker is 0.33 ng/mL (S/N=18), while the LOD of 1.2 kb RNA marker is 0.081ng/mL (S/N=4). The detector has approximately 3 orders of dynamic range.



**Figure 8.** a) Linearity of detection response using LIF detector, plotting peak area versus concentration over the range of 0.33 ng/mL to 333 ng/mL of the 1.2 kb RNA Marker b) log (detection response) to log (concentration in ng/mL) plot to have the lower concentration part of the linearity plot visible.

p6

[Download this technical note](#)

### Conclusions

In conclusion, a CGE-LIF method was optimized to resolve RNA from 0.2 kb to 6.5 kb in size. The optimized method demonstrates good resolving power, MT and CPA% repeatability, excellent linearity from 0.33 ng/mL to 333 ng/mL for 1.2 kb RNA Marker, LOQ as low as 0.33 ng/mL and provides the flexibility of further method modification.

### Reference

1. Julia Khandurina, Hur-Song Chang, Bart Wanders, and Andras Guttman, *AN ULTRAFAST METHOD TO EVALUATE RNA QUALITY*

**Who is SCIEX?** SCIEX company's global leadership and world-class service and support in the capillary electrophoresis and liquid chromatography-mass spectrometry industry have made it a trusted partner to thousands of the scientists and lab analysts worldwide who are focused on basic research, drug discovery and development, food and environmental testing, forensics and clinical research.

**Contact Us:** [sciex.com/contact-us](http://sciex.com/contact-us)

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

© 2018 AB SCIEX. SCIEX is part of AB SCIEX. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

Publication number: RUO-MKT-02-8017-B 08/2018



**Headquarters**  
500 Old Connecticut Path, Framingham, MA 01701, USA  
Phone 508-383-7800  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





## RNA Analysis for CRISPR by Capillary Electrophoresis with Laser-Induced Fluorescence Detection

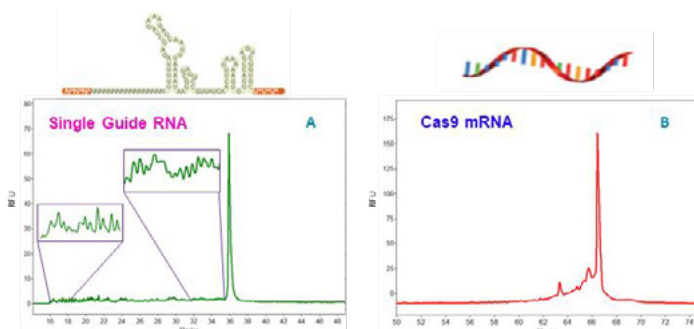
### CE-LIF

Jane Luo, Tingting Li, Marcia Santos, Fang Wang, Sahana Mollah and Handy Yowanto  
SCIEX Separations, Brea, California

### Introduction

CRISPR (clustered regularly interspaced short palindromic repeats) is an immunological defense mechanism employed by bacteria against invading pathogens. Cas9 (CRISPR associated protein 9) is an RNA-guided DNA endonuclease. Due to its ability to perform site directed DNA cleavage and to trigger homologous recombination, Cas9 has been utilized as a gene editing tool to introduce gene inactivation and genome modification. In the pharmaceutical industry, CRISPR can improve existing therapies, help find drug targets and test drug candidates.<sup>1</sup>

Two RNA molecules are essential for CRISPR: single guide RNA (sgRNA) and Cas9 mRNA. During the development of RNA therapeutics, it is often advantageous to co-deliver the sgRNA and the Cas9 mRNA. Purity analysis of these two molecules currently requires two methods: denaturing agarose gel electrophoresis for Cas9 mRNA and denaturing polyacrylamide gel electrophoresis for sgRNA. In this technical note, we describe a CE-LIF method that can analyze both sgRNA and Cas9 mRNA using the same gel matrix. Purity information on both molecules can be obtained in a single sequence run, leading to not only a simplified analysis process, but also savings on time and resources.



**Figure 2.** Analysis of HPRT single guide (sg) RNA (A) and cas9 mRNA (B) by CE-LIF. The full length HPRT sgRNA was detected as a single peak around 36 minutes when run in 100-R gel at 30kV for 50 minutes. Oligonucleotides that failed one or multiple coupling during oligo synthesis are located on the left side of the main peak, from 16 to 35.8 minutes with some zoomed-in areas shown in insets. The cas9 mRNA sample was run in diluted 100-R gel for 75 minutes at 6kV. Sample peaks include a major peak at 66.7 minutes plus a few minor peaks between 63 to 66 minutes. The raised baseline indicates the presence of many additional degradation products with gradually decreasing sizes.



**Figure 1.** The PA 800 Plus Pharmaceutical Analysis System (A) and the ssDNA 100-R Kit (B).

### Key Features

- Same gel and capillary for the analysis of sgRNA and cas9 mRNA for CRISPR
- Capability to analyze small and large RNA in one sequence
- Single base resolution for oligos down to 15 nucleotides in length with gel matrix in the ssDNA 100-R kit
- Capability of RNA size estimation in the range of 0.2 kb to 10 kb with diluted gel matrix, aiding in understanding of RNA degradation
- Good repeatability with %RSD of MT below 0.7% and %RSD of percent corrected peak area below 2.8%

## Materials, Instrument and Methods

**Materials:** The eCAP ssDNA 100-R kit (PN 477480, Figure 1B) and LIF Performance Test Mix (PN: 726022) were from SCIEX, Framingham, MA. The Human HPRT Single Guide RNA (sgRNA) (100 nt, PN 217541712) was obtained from IDT, Coralville, Iowa. CleanCap Cas9 mRNA with modification (PN L-7206-20, 20 µg) was from TriLink BioTechnologies, San Diego, California. Urea (PN 29700), Nuclease free water (PN AM9932) and SYBR Green II RNA gel stain, 10,000x concentrate in DMSO (PN S7564) were obtained from Thermo Fisher Scientific, Waltham, MA. Polyvinyl-pyrrolidone (PVP, PN 437190), Transcript RNA markers 0.2-10kb (PN R7020) and 10xTBE buffer (PN 574795), Molecular Biology Grade were from Sigma-Aldrich, St. Louis, MO. The 0.2 µm syringe filter (PN 4602) was from PALL Corporation, Port Washington New York. Rainin LTS filter tips were from Mettler Toledo, Oakland, California.

**Instrument and software:** A PA 800 Plus Pharmaceutical Analysis System (Figure 1A) equipped with LIF detector and solid-state laser with excitation wavelength at 488 nm and a 520 nm band pass emission filter were from SCIEX, Framingham, MA. Data acquisition and analysis were performed using 32 Karat software V10.

**Tris-Borate-7 M Urea Buffer Preparation:** This step must be done one or a few days before running samples. To rehydrate the buffer, 135 ml of 0.2 µm filtered deionized water was added to the bottle containing dry tris-borate in the ssDNA 100-R kit. The solution was mixed with a clean stirring bar for about 30 min or until boric acid was completely dissolved. Then, the entire amount of urea from the urea bottle in the ssDNA 100-R kit was added while the solution was being stirred. The urea was completely dissolved after about 2 hours and the solution became clear. This buffer should be good for one month if stored at 2°C to 8°C. Before sample run, the required amount of buffer was removed and filtered through a 0.2 µm filter. SYBR Green II dye was added at a 1 to 50,000 dilution. About 3.5 ml of this dye-containing buffer was used for each set of 6 injections.

**Preparation of 100-R gel:** Five milliliters of freshly filtered Tris-Borate- 7 M Urea buffer was added to the lyophilized gel. The solution was stirred for about 3 to 6 hours until the gel is completely dissolved. Alternatively, the gel solution was placed on a LABQUAKE rotator (Barnstead International, Dubuque, IOWA) in a cold room (2°C to 8°C) for 72 hours with gentle rotation. For analysis of single guide RNA (100 nt) and pd(A) 40-60 Test Mix, undiluted 100-R gel was used as is. For analysis of cas9 mRNA and large RNA molecules, 536 µl of the 100-R gel was diluted 2.8 fold to 1500 µl with filtered Tris-Borate-7 M Urea

buffer. After thorough mixing by gentle pipetting and swirling, the diluted gel was spun on a Qualitron DW-41 micro centrifuge at 2000 x g (6400 rpm) for 4 min to remove bubbles.

**Cartridge Assembly:** DNA capillary (PN 477477) was installed per instructions on kit insert (PN 726479) in the ssDNA 100-R kit. The total capillary length was 40.2 cm with 30 cm as the effective length.

**LIF Calibration:** To ensure consistent response of LIF detector throughout this study, the LIF detector was calibrated using LIF Calibration Wizard and Performance Test Mix (PN: 726022), following the instructions in LIFluor Enhance user's guide (PN 725824). This calibration was done whenever capillary or LIF detector or the laser was changed.

**Preparation of Buffer Trays and Sample Trays:** Vial positions for buffer trays are indicated in Figure 3 with rows 1 to 3 for tests with guide RNA and pd(A) 40-60 Test Mix and rows 4 to 6 for tests with Cas9 mRNA and larger RNA molecules. All solutions were pipetted with filter tips. Each "NF Water" vials were filled with 1.5 ml nuclease free (NF) water. Waste vial was filled with 1 ml NF water. "Dilut. 100R" vial was filled with 1.5 ml diluted 100-R gel. "100R Gel" vial was set up with a microvial containing 210 µl 100-R gel. "Sep Buffer" vials were filled with 1.6 ml filtered Tris-Borate-7 M Urea buffer containing SYBR Green II dye.

**Sample Preparation:** Filter tips were used for all steps. Lyophilized sgRNA was resuspended in 10 mM Tris, pH 7.5, 0.1 mM EDTA to a concentration of 100 µM or 3.246 µg/µl and stored at -80°C in 10 µl aliquots. Before use, it was then diluted with nuclease free water to 25 ng/µl and treated at 80°C for 2 minutes. Immediately after heat treatment, it was placed on ice for 5 minutes before being transferred to a microvial (PN 144709) and loaded onto the sample inlet tray. The pd(A) 40-60 Test Mix from the 100-R kit was resuspended in 500 µl of nuclease free water and stored at -20°C in 50 µl aliquots. It was further diluted 2 fold with NF water before injection. The Sigma Transcript RNA markers and Cas9 mRNA were aliquoted upon first thawing and stored at -80°C in aliquots. Before use, they were diluted to 25 ng/µl (marker) and 10 ng/µl (Cas9 mRNA), and treated at 65°C for 5 minutes, followed by 5 minutes on ice before transferred to a microvial and loaded onto the instrument.

**Instrument Set up:** Conditions used for analysis of sgRNA and pd(A) 40-60 Test Mix were as described in Figures 4-8. Separation was done at 30 kv for 50 minutes for sgRNA and at 12.4 kv for 50 minutes for the pd(A) 40-60 Test Mix. Conditions for analysis of Cas9 mRNA and the Transcript RNA markers were carried out as shown in Figures 4 to 6 and 8-9.

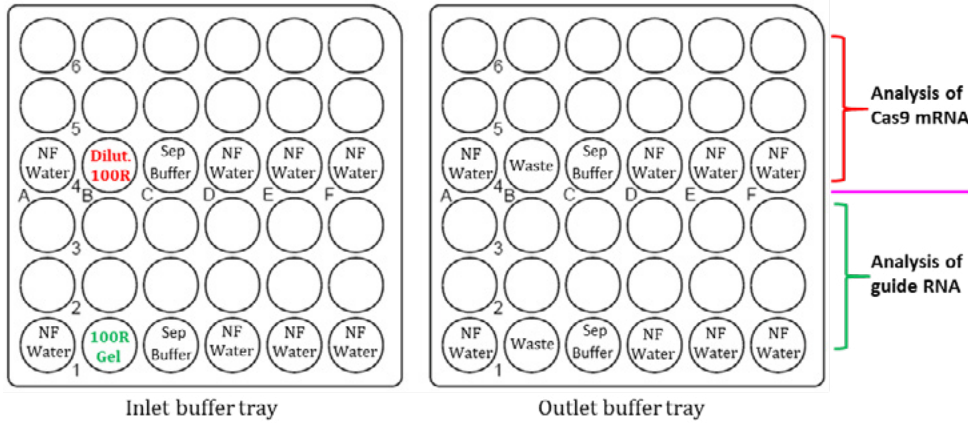


Figure 3. Schematic set up for Inlet (left) and Outlet (right) buffer trays for analysis of guide RNA (rows 1-3) and Cas9 mRNA (rows 4-6). Each row is enough for at least 6 sample injections. Additional rows can be set up if more samples need to be analyzed.

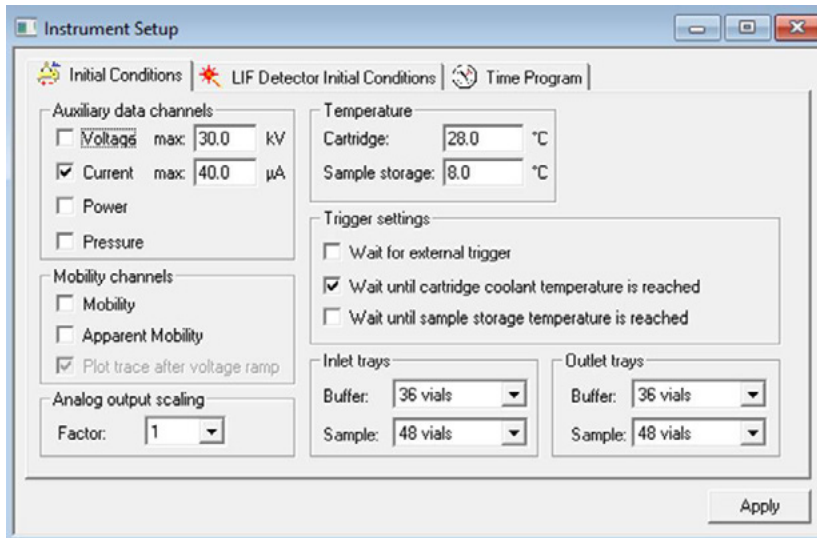


Figure 4. Settings for “Initial Conditions” tab. Numbers shown are for Cas9 mRNA analysis. Same conditions were used for analysis of sgRNA and pd(A) 40-60 except that the cartridge temperature was set to 20°C and 30°C for sgRNA and pd(A) 40-60 respectively.

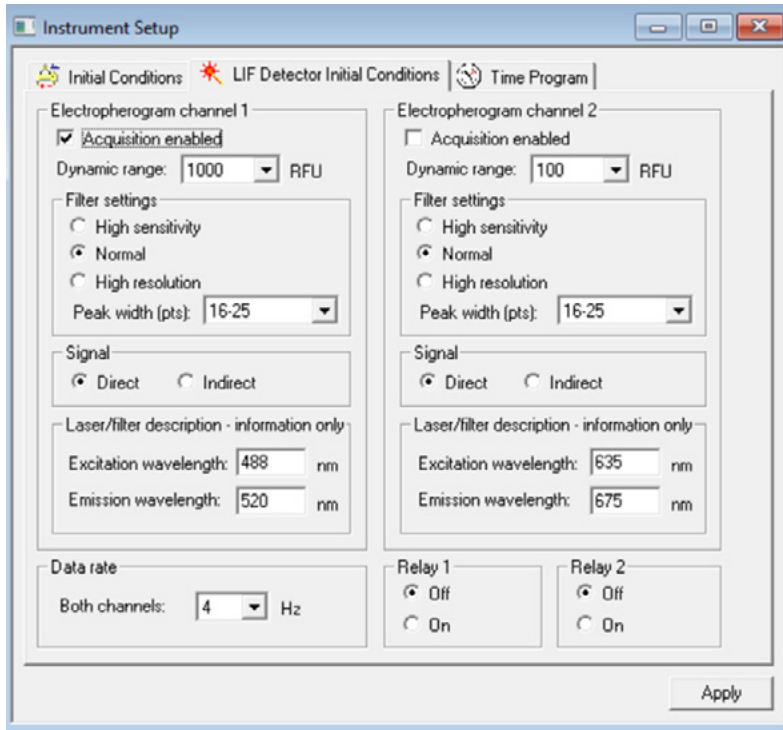


Figure 5. Settings for “LIF Detector Initial Conditions” tab for all RNA analysis.

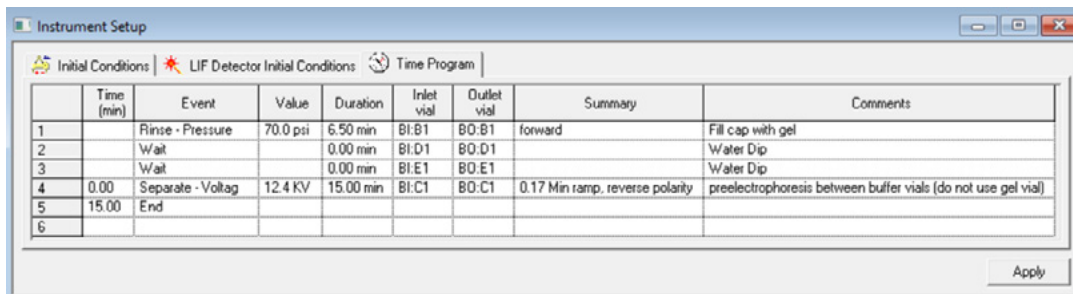


Figure 6. Time Program Settings for the Conditioning Method. Settings shown are for analysis of sgRNA and pd(A) 40-60. Same conditions were used for Cas9 mRNA analysis except that vial positions were adjusted to row 4.

	Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1		Wait		0.00 min	BI:A1	BO:A1		H2O dip
2		Inject - Voltage	4.0 KV	2.0 sec	SI:A1	BO:C1	Override, reverse polarity	Sample injection against buffer vial
3		Wait		0.00 min	BI:F1	BO:F1		H2O dip
4	0.00	Separate - Voltag	30.0 KV	50.00 min	BI:C1	BO:C1	0.17 Min ramp, reverse polarity	Use buffer vials
5	1.00	Autozero						
6	50.00	End						
7								

**Figure 7.** Time Program Settings for the Separation Method of sgRNA. The separation voltage was changed to 12.5 kv for analysis of pd(A) 40-60 test mix or smaller sized RNA at 30 to 70 nt length.

	Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1		Rinse - Pressure	20.0 psi	3.00 min	BI:B1	BO:B1	forward	Filling with 100-R gel
2		Wait		0.00 min	BI:A1	BO:A1		ddH2O dip
3		Laser - Off						
4	3.01	End						
5								

**Figure 8.** Time Program Settings for the Shut Down Method for analysis of sgRNA or smaller RNA. The vial positions were changed to row 4 for analysis of Cas 9 mRNA and large RNA.

	Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1		Wait		0.00 min	BI:A4	BO:A4		Water Dip
2		Wait		0.00 min	BI:F4	BO:F4		Water Dip
3		Inject - Voltage	1.0 KV	2.0 sec	SI:A1	BO:C4	Override, reverse polarity	sample injection against buffer vial
4	0.00	Separate - Voltag	6.0 KV	75.00 min	BI:C4	BO:C4	0.17 Min ramp, reverse polarity	Separation
5	1.00	Autozero						
6	75.00	End						
7								

**Figure 9.** Time Program Settings for the Separation Method of Cas9 mRNA or large RNA.



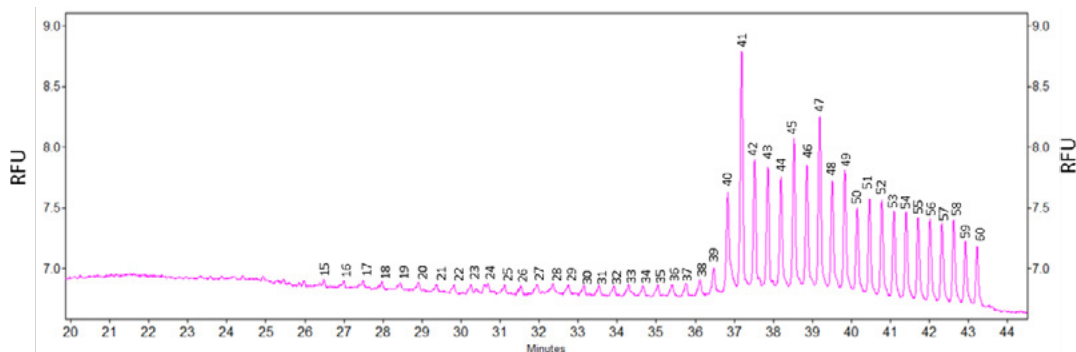
## Results and Discussions

**Analysis of sgRNA:** In CRISPR-Cas9 workflow, the guide RNA provides sequence-specificity and targets the Cas9 nuclease to the site of gene editing where the Cas9 nuclease performs double stranded DNA cleavage. The guide RNA occurs naturally as a two-molecule complex consisting of a target-specific crRNA (CRISPR RNA) bound to a tracrRNA (trans-activating crRNA) that directs the binding of RNAs to Cas9 nuclease. Recently, the use of sgRNA has become very popular in pharmaceutical industry. A sgRNA is a single RNA molecule designed to contain the sequences of both crRNA and tracrRNA. The diagram on top of Figure 2A shows the sequence and secondary structure of the sgRNA molecule used in this study. It was designed and synthesized by IDT with a brand name of Alt-R CRISPR-Cas9 sgRNA.<sup>2</sup> The target gene is human HPRT gene (Hypoxanthine-guanine phosphoribosyl-transferase). Nucleotides shown in bold white are 2'OMe bases; the asterisks indicate phosphorothioate linkages. Chemical modifications on Alt-R CRISPR-Cas9 sgRNAs increase their stability, potency, and resistance against nuclease activity. After synthesis, this sgRNA was desalted and dried without further purification. Figure 2A shows the results obtained with this sgRNA using ssDNA 100-R gel and Tris-Borate-7M Urea buffer containing SYBR Green II dye. The HPRT sgRNA was diluted with nuclease free water to 25 ng/ $\mu$ l and analyzed by CE-LIF. It was detected as a main peak around 36 minutes when electrophoresed at 30kv for 50 minutes. As expected for desalted oligonucleotides, many small peaks representing the oligonucleotide species that failed one or

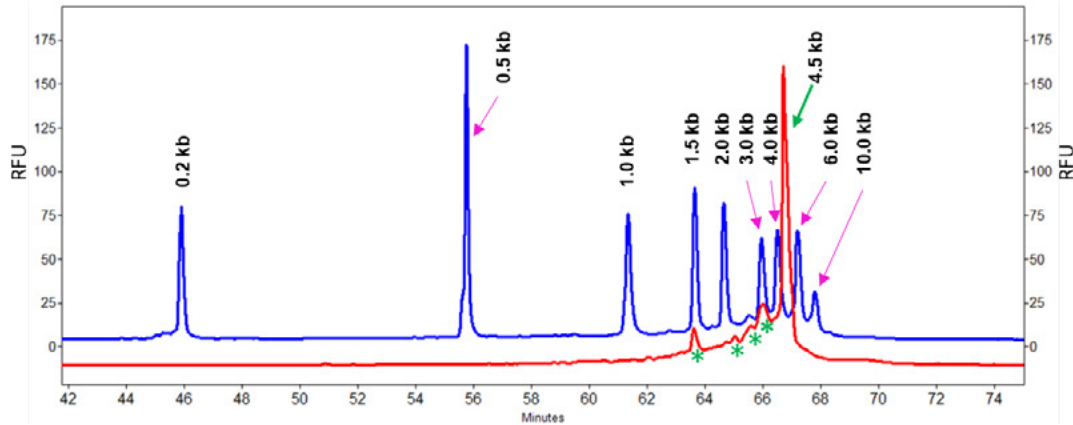
multiple coupling cycles during oligonucleotide synthesis were detected on the left side of the main peak, from 16 to 35.8 minutes. Some areas in this region were displayed in insets as zoomed-in images to show how these small peaks look like. After calculating the corrected peak area for all impurity peaks and the main peak, the purity of this sgRNA was determined as 45.07%. This data indicates that this sgRNA has the typical purity level expected for a longer oligonucleotide.<sup>3</sup>

**Analysis of pd(A) 40 to 60 Test Mix:** Although the main purpose of this technical note is to analyze sgRNA and Cas9 mRNA, some CRISPR customers may use naturally occurring crRNA and tracrRNA with sizes around 30 to 60 nt. Therefore, the feasibility of using the eCAP ssDNA 100-R kit for the analysis of crRNA and tracrRNA was tested by running pd(A) 40 to 60 Test Mix. As shown in Figure 10, all expected 21 peaks with sizes from 40 nt to 60 nt were detected. In addition, n-1 peaks with sizes smaller than 40 nt were detected too. Single base resolution was achieved from 15 nt to 60 nt. The high resolving power of the 100-R gel can also show the appearance of shoulders on major peaks such as on the 23nt and 24 nt peaks. These shoulders maybe the dephosphorylated form of each oligonucleotide or other impurities.

**Detection of RNA degradation in Cas9 mRNA:** Since the Cas9 mRNA is used to express the Cas9 nuclease, it is critical to test its purity and integrity before it is co-delivered with sgRNA to targeted cells. A Cas9 mRNA used in this study expresses a version of the *Streptococcus pyogenes* SF370 Cas9 protein. It contains a naturally occurring Cap 1 structure, polyadenylated and with a modified uridine to mimic a fully processed mature mRNA. Results in Figure 2B and Figure 11 indicated that



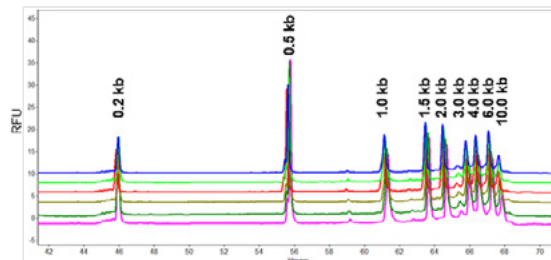
**Figure 10:** Analysis of pd(A) 40 to 60 Test Mix with 100-R gel using CE-LIF. The 21 expected single stranded DNA oligonucleotides with sizes from 40 nt to 60 nt were detected. In addition, the n-1 peaks with sizes smaller than 40 nt were also detected.



**Figure 11.** Detection of degradation in Cas9 mRNA sample. Cas9 mRNA sample (10 ng/μl) was separated in 100-R gel diluted 2.8 fold using the Tris-Borate-7 M Urea buffer. The green arrow indicates the intact Cas9 mRNA which is 4.5 kb. RNA degradation was detected as indicated by the green asterisks. Based on the overlay of the Cas9 mRNA trace (red) and the trace obtained with RNA markers (blue), significant RNA degradation species have sizes around 3 kb, 2.5 kb, 2 kb and 1.5 kb (green asterisks).

Cas9 mRNA was detected as a main peak around 66.7 minutes with an estimated size of 4.5 kb, consistent with the size provided by the manufacturer. RNA degradation was observed in this Cas9 mRNA sample. Corrected peak area percent (%CPA) for the main peak is 58.08%. Through comparison of electropherograms obtained with Cas9 mRNA and with the Transcript RNA markers, it was estimated that the main degradation species have sizes as 3 kb, 2.5 kb, 2 kb and 1.5 kb with %CPA of 16.83%, 3.70%, 4.65% and 7.22% respectively. Impact of this degradation will be investigated in a separate study. In addition, the raised baseline indicates that additional degradation species with gradually decreasing sizes were also present. Since the fragments in the Transcript RNA markers have different sequences than the Cas9 mRNA, the size estimation for the RNA degradation species based on RNA markers can be slightly different from their true theoretical sizes. In addition, since mRNA is single stranded, the impact of sequence composition on size estimation is more significant than with double stranded DNA size analysis. However, the estimated size values for the degradation species can be valuable in understanding Cas9 mRNA sample degradation.

**Method Repeatability:** The Transcript RNA markers with sizes ranging from 0.2 kb to 10 kb was injected 6 times to test the method repeatability using the diluted 100-R gel. As shown in



**Figure 12.** Overlay of six consecutive runs of Sigma Transcript RNA markers with sizes from 0.2 kb to 10 kb with diluted 100-R gel.

**Table 1.** The %RSD of migration time and corrected peak area percentage (CPA%) for each of the 9 RNA fragments in RNA ladder was analyzed in the repeatability study.

RNA Ladder Size (in bases)	%RSD for Migration Time	%RSD for CPA%
200	0.32	1.92
500	0.64	1.16
1000	0.37	1.77
1500	0.35	1.83
2000	0.34	1.34
3000	0.36	2.71
4000	0.34	1.50
6000	0.33	1.92
10000	0.33	1.39

Table 1 and Figure 12, excellent repeatability was obtained. The relative standard deviation percent (%RSD) for migration time was lower than 0.7% for all fragments. The %RSD for corrected peak area percentage was lower than 2.8% for all fragments.

## Conclusions

- Analysis of both sgRNA and Cas9 mRNA using the same 100-R gel at different concentrations provides convenience to users who need to analyze both of them on a routine basis
- The eCAP ssDNA 100-R kit supplies pre-formulated gels to save time for users
- Single base resolution down to 15 nt provides confidence in analyzing n-1 species in short oligos such as crRNA and tracrRNA
- Capability of estimating impurity size provides clues for understanding RNA degradation in Cas9 mRNA
- Excellent method repeatability ensures reliable results

## References

1. Three ways CRISPR will change the drug world. (2019) Pharmaceutical industry finding novel uses for gene editing tools. <https://www.optum.com/resources/library/crispr.html>.
2. Alt-R CRISPR-Cas9 sgRNAs. Chemically synthesized and modified single guide RNAs for outstanding CRISPR performance and quality. IDT product information brochure. CRS-10142-FL 07/19.
3. Capillary electrophoresis of oligonucleotides. (2011) Integrated DNA Technologies (IDT).

## Acknowledgements

We thank Mukesh Malik, Peter Holper and Zuzana Demianova for helpful discussions.

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. RUO-MKT-02-10432-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



## Highly Precise DNA Sizing on the SCIEX Genetic Analysis System

### DNA Fragment Analysis

Keith Roby, Mark Dobbs, Doni Clark, Scott Boyer, and Graham Threadgill  
*Beckman Coulter, Inc.*

### Introduction

Capillary electrophoresis is rapidly supplanting conventional slab gel electrophoresis in moderate- to high-throughput laboratories that desire automated separations with a simple user interface. Capillary systems offer a level of automation and ease of use that cannot be achieved with slab gels. There is no ambiguity as to the identity of the sample in a capillary system, and the use of internal size standards allows a level of precision that far exceeds the visual comparison of adjacent lanes on a slab gel. Our data demonstrates that the SCIEX CEQ/GeXP Genetic Analysis System precisely sized microsatellite amplification products in the size range of 200, 300, and 400, to a standard deviation of  $\leq 0.27$  bases, regardless of which capillary within an array, run, gel lot, or instrument was used to separate the sample.

### Materials and Methods

Three different loci from the Weber panel were used in this study:

Locus	D18S858	D6S1040	D14S742
Marker Name	ATA23G05	GATA23F08	GATA74E02
Allele Size Range (nt)	193–208	257–285	395–415
Expected Alleles for Genomic DNA 1331-01	196, 196	277, 281	403, 403
Dye/Trace Color	D3/Green	D4/Blue	D2/Black

Primer sets were obtained from Research Genetics, Inc. (Huntsville, Alabama). The genomic DNA was CEPH Utah pedigree 1331, repository DNA sample NA07057A obtained from the Coriell Institute for Medical Research (Camden, NJ). Labeled primers were tested for quality by dilution to 30 fM in deionized formamide, and separated on the GeXP. Primer pairs were quantitated using a DU® 7500 UV Spectrophotometer. Genomic DNA was quantitated on a FLUOstar-P Microplate

Fluorometer (BMG Lab Technologies, Offenburg, Germany) using Pico Green (P-7581; Molecular Probes, Eugene, Oregon). Lambda DNA was used as a quantitative standard for this assay. Individual amplification reactions were mixed as follows:

10X AmpliTaq Gold PCR Buffer	1.0 $\mu$ L
5X dNTPs (1.25 mM)	0.8 $\mu$ L
Primer Mix (2.5 mM each)	0.5 $\mu$ L
MgCl <sub>2</sub> (25 mM)	0.6 $\mu$ L
DNA (20 ng/ $\mu$ L)	2.0 $\mu$ L
Enzyme (5U/ $\mu$ L)	0.05 $\mu$ L
Water	5.05 $\mu$ L
<b>Total Reaction Volume</b>	<b>10.0 mL</b>

Cycling conditions were: 95° C—10 minutes, 1 cycle to activate the enzyme (AmpliTaq Gold,\* Perkin Elmer, Branchburg, NJ), 95° C—1 minute, 57° C—1 minute, 72° C—1 minute, 40 cycles (MJ PTC 200 Thermocycler, MJ Research, Inc. Waltham, MA). The individual samples were then run on a GeXP to determine the dilutions needed to balance the signal intensities of the size standards and amplification products. Mixtures large enough for 96 wells were prepared, and 40  $\mu$ L of mixture was aliquoted into each well. For each sample well, the equivalent of 0.5  $\mu$ L of the three PCR amplifications, 0.5  $\mu$ L of the 608098 DNA Size Standard Kit 400 bp, and 39  $\mu$ L of deionized formamide were added. Samples were run on a GeXP under the conditions given in Table 1.

Seven full plates were electrophoresed with two of three variables held constant (Table 2). The three variables were gel lot, array, and instrument. During the course of these experiments, several independent preparations of sample were done, as indicated in Table 3. In two cases, the same plate was used for a second injection.

p1

[Download this technical note](#)



## Results and Discussion

The alleles generated in this experiment are illustrated in Figure 1. In all four cases, a +A artifact peak was generated, representing a different fraction of the overall locus-specific amplification product. The true allele, not the +A form of the allele, was used in calculating the mean and standard deviation of the allele sizes in Table 4. Note that, in this table, two of the data sets are duplicated as indicated in the Experiment column.

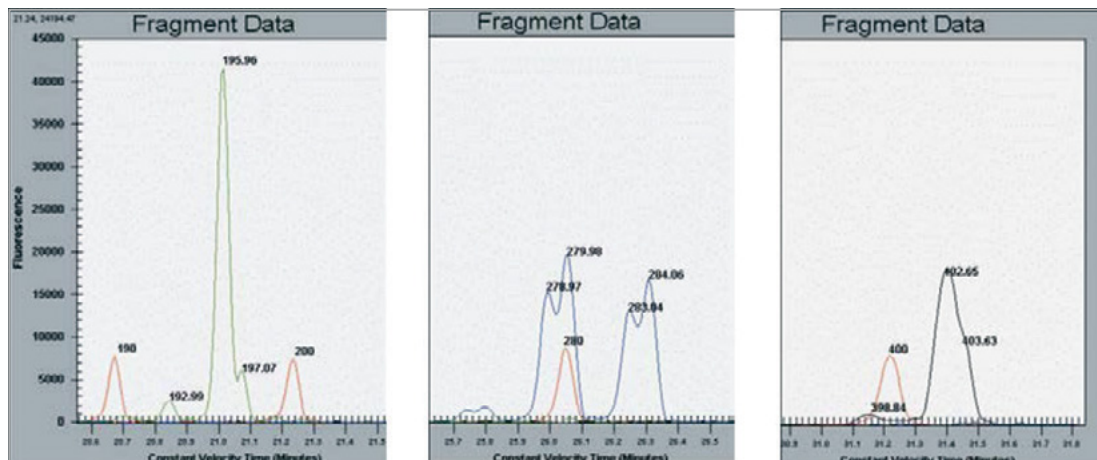
The four peaks differed slightly from their expected sizes with the differences ranging from 0.1 to 1.7 bases. The differences were consistent across all experiments and appear to be independent of fragment length. Such variations are likely due to slight differences in composition of size standards, gel formulations, run temperatures, and other differences in the electrophoresis platforms. When comparing data obtained from different fragment sizing platforms, a series of known samples is often used to determine the corrections that must be applied in order to obtain the correct allele assignments. Due to the high level of precision demonstrated above, the size corrections determined at any locus on the GeXP alone should be constant as long as the fragments are run using the same temperature and voltage.

Capillary Temperature	50°C
Denature Temperature and Time	90°C for 120 seconds
Injection Voltage and Time	2.0 kV for 30 seconds
Separation Voltage and Time	6.0 kV for 35 minutes

**Table 1.** Frag-3 Run Method.

Experiment	Gel	Array	Instrument
1	1	1	1
2	1	1	2
3	1	1	3
4	1	2	3
5	1	3	3
6	2	3	3
7	3	3	3

**Table 2.** Plates Run with Two of Three Variables Held Constant.



**Figure 1.** The four fragments used in the accuracy and precision tests. In each case, the first peak of the doublet was selected. The second peak was regarded as the +A artifact. The black peak includes the +A peak as shoulder on the main peak. The size standards are the peaks in red.

Experiment	Amplification	Mixture	Injection	Operator
1	1	1	1st	1
2	1	1	2nd	1
3	2	1	1st	2
4	2	2	1st	2
5	3	1	1st	2
6	3	2	1st	2
7	3	2	2nd	2

**Table 3.** Independent Amplifications and Preparation of Sample Mixtures for the Accuracy and Precision Tests.

Within each of the seven plates examined, there was no noticeable drift of sizes from the first to the last row of a plate and no rows or capillaries that demonstrated unusual deviation from the mean sizes calculated from the fragment lists. Thus, samples appear to be stable within the time it takes to run an entire plate (approximately nine hours for 400 bases of separation). The capillary array that was used on three different instruments, consisting of more than one third of its expected useable lifetime, did not demonstrate a trend toward producing progressively higher or lower called sizes with increasing run numbers. Finally, the three different gel lots tested gave the same results, indicating a high level of reproducibility in the process used to prepare the Beckman Coulter polymer mixture. As expected, independent amplifications and preparations of the sample mixtures by different operators had no effect on the results. A second injection of sample was performed twice in the above experiments with no ill effects. While there is negligible sample depletion during each injection, some signal loss is expected due to formamide degradation, a process that is accelerated at high temperatures.

As a whole, the standard deviation observed with the 196-base fragment was less than half of the value obtained for the other three fragments. The reasons for this high level of precision are not known at this time. It is possible that the deviation from the mean fragment length is smaller due to the short residence time in the capillary, allowing less time for diffusion. It is also possible that, under the run conditions used here, the population of 196-base fragments is more homogeneous in shape than the other three populations due to its particular sequence. Clearly, more examples would be required before any conclusions could be drawn.

In general, the ability to properly determine the correct number of repeat units in STR amplification products requires that a system be more precise than half the length of a repeat unit over a reasonable range of expected sizes. For example, to assign a tetranucleotide repeat allele to its proper size bin (range of observed sizes for a unique allele), a precision of better than two bases is required, and onebase precision is required for dinucleotide repeats. To be of general utility in large gene-mapping projects, however, a fragment sizing system must reproducibly determine fragment lengths of alleles containing imperfect repeats. Thus, better than 0.5-base precision is desirable since alleles containing repeat units of any size may, in rare cases, differ in overall length by only a single nucleotide. We demonstrated that, under a variety of circumstances likely to be encountered by GeXP users, the standard deviation about the mean is no greater than 0.26 bases.

## Conclusions

Among the simple tandem repeat fragments tested, the GeXP demonstrates outstanding reproducibility in calling fragment sizes. This level of precision would facilitate properly identifying alleles that differ by only one nucleotide in length up to at least 400 bases. The precision is achieved with consistency in the manufacturing of gels, the coating of capillaries, the uniformity of capillary temperature and run voltage, and the ability of the software to accurately identify peak centers.

Experiment		Expected Allele> Size (bases)	196	277	281	403
		Observed Sizes				
1	Instrument 1	Mean	196.19	278.59	282.62	402.45
		S.D.	0.05	0.12	0.10	0.11
2	Instrument 2	Mean	196.21	278.86	282.94	402.69
		S.D.	0.05	0.13	0.13	0.14
3	Instrument 3	Mean	196.19	278.40	282.48	402.28
		S.D.	0.03	0.13	0.13	0.09
Across Instruments		Mean	196.19	278.59	282.68	402.47
		S.D.	0.05	0.12	0.24	0.21
3	Array 1	Mean	196.19	278.40	282.48	402.28
		S.D.	0.03	0.13	0.13	0.09
4	Array 2	Mean	196.14	278.67	282.74	402.49
		S.D.	0.05	0.13	0.12	0.11
5	Array 3	Mean	196.07	278.69	282.77	402.45
		S.D.	0.05	0.12	0.14	0.13
Across Arrays		Mean	196.13	278.58	282.66	402.40
		S.D.	0.18	0.19	0.19	0.15
5	Gel 1	Mean	196.07	278.69	282.77	402.45
		S.D.	0.06	0.12	0.12	0.13
6	Gel 2	Mean	195.95	278.46	282.53	401.98
		S.D.	0.03	0.09	0.09	0.10
7	Gel 3	Mean	195.91	278.50	282.57	401.95
		S.D.	0.03	0.06	0.06	0.11
Across Gels		Mean	195.98	278.54	282.62	402.13
		S.D.	0.08	0.15	0.15	0.27
Across All Variables		Mean	196.10	278.57	282.65	402.33
		S.D.	0.12	0.20	0.20	0.26

**Table 4.** Observed Means and Standard Deviations from 96-Well Plates.





# Plasmid analysis

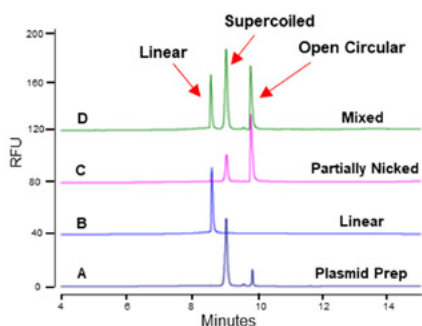
## Simple, Fast and Robust Method for Plasmid Purity Testing and Degradation Monitoring

### Capillary Electrophoresis-Laser Induced Fluorescence Detection (CE-LIF)

Jane Luo, Tingting Li, and Mukesh Malik  
SCIEX, Brea, California

#### Introduction

Plasmid DNA is often used as the gene delivery vehicle for vaccinations and gene-therapy products. It is also an important intermediate in making Adeno-associated viruses and lentiviruses. The Food and Drug Administration (FDA) guidelines recommend manufacturers include a specification of > 80% supercoiled plasmid content for bulk release (Ref 1). Capillary electrophoresis with laser-induced fluorescence detection (CE-LIF) provides a rapid, sensitive, reproducible and automated method for the quantitative analysis of plasmid DNA isoforms. Advantages of CE-LIF over slab-gel electrophoresis include rapid analysis time, minimal sample handling, simpler set up and preparation time, greater sensitivity and reproducibility, more accurate quantitation, automated sample loading, and reduced waste generation. This technical note describes a CE-LIF method for plasmid analysis. In this method, the supercoiled, linearized and open circular plasmid isoforms for 5 kilobase and 7 kilobase plasmids can be reproducibly separated with baseline resolution in less than 15 minutes.



**Figure 2.** Confirmation of different isoforms in an untreated 5 kb plasmid preparation (A) using linearized (B) and nicked plasmids (C) as controls. The major peak in the untreated Plasmid Preparation (A) is the supercoiled isoform; the smaller peak is the open circular isoform. In the "Mixed" sample (D), the first peak is the linear isoform, the second peak is supercoiled isoform, and the third one is the open circular isoform.



**Figure 1.** The PA 800 Plus Pharmaceutical Analysis System (A), the dsDNA 1000 Kit (B) and the LIFluor Enhance Stain (C).

#### Key Features

- Complete workflow solution for plasmid analysis including hardware, software and reagent kits
- Baseline resolution of supercoiled, linear and open circular plasmid isoforms
- Sample separation takes less than 12 minutes
- Increased throughput for analysis via automated sample separation and data analysis
- Robust, good repeatability quantitative assay with RSD of 0.23% for migration time and 0.61% for % area
- Capable of monitoring low-level plasmid degradants
- Less waste generation than slab gel method
- LIFluor Enhance stain in this method is safer than ethidium bromide used in slab gel method



## Materials and Instrument

### Materials:

The eCAP dsDNA 1000 kit (PN 477410, Figure 1B), the LIFluor Enhance fluorescent stain (PN 477409, Figure 1C), dsDNA 1000 Test Mix (PN: 477414) and LIF Performance Test Mix (PN: 726022) were from SCIEX, Framingham, MA. A 10X concentrated stock solution of Tris Borate EDTA (10x TBE, PN T4323) was from Sigma Aldrich, St. Louis, MO. The 5 kilobase (kb) Plasmid DNA (PN E1541) was from Promega, Madison, Wisconsin. The 7 kb Plasmid DNA was from an internal project at SCIEX, Framingham, MA. The nicking enzyme, Nb.BsmI (PN R0706) and BamHI-HF restriction enzyme (PN R3136) were from New England BioLabs, Boston, Massachusetts.

**Instrument and software:** A PA 800 Plus Pharmaceutical Analysis System (Figure 1A) equipped with LIF detection and solid-state laser with excitation wavelength at 488 nm and a 520 nm band pass emission filter were from SCIEX, Framingham, MA. Data acquisition and analysis were performed using 32 Karat software V10.

## Methods

**Gel buffer reconstitution:** This step must be done one day before running samples. To rehydrate the gel buffer, 20 ml of 0.2 µm filtered deionized water was added to the gel buffer vial. After the gel buffer vial was securely capped, the vial was placed on a rotator or the gel mixture was gently stirred with a small stirring bar for up to 24 hours or till the dried gel was completely dissolved. The hydrated gel should be good for one month if stored at 4°C. Then, the gel solution was diluted 1:10 with 1X TBE (90 mM Tris, 90 mM Borate, and 2 mM EDTA pH 8.3). The diluted gel was filtered through a 0.45 µm filter. 1x TBE was prepared by a 10x dilution of TBE stock solution with deionized water followed by passing through a 0.2 µm filter. LIFluor Enhance stain was added to the gel at a concentration of 5 µl stain in 6 ml gel (Ref 2).

**Cartridge Assembly:** DNA capillary (PN 477477) was installed per instructions on kit insert (PN 726412) in the dsDNA 1000 kit. The total capillary length was 40.2 cm with 30 cm as the length to the detection window. Since the inner wall of the DNA capillary is coated, the cartridge assembly should be carried out in a timely manner. Excessive exposure to air may damage the inner coating and also cause clogging. The capillary ends must be immersed in liquid (water or buffer) as soon as the cartridge assembly is complete to prevent the coating from drying out.

**Sample Preparation:** In order to identify different plasmid isoforms, linearized and open-circular plasmids were prepared as controls. To prepare the linearized isoform, a single-cutting

restriction endonuclease, BamHI-HF, was used to digest plasmids following manufacturer's instructions. To obtain the open circular isoform, plasmid preparations were either partially or fully digested with a nicking enzyme according to the manufacturer's instructions. All plasmid samples were diluted with deionized water to 10 ng/µl for analysis on the PA 800 Plus instrument.

**LIF Calibration:** To ensure consistent response of LIF detector throughout this study, the LIF detector was calibrated using LIF Calibration Wizard and Performance Test Mix (PN: 726022) following the instructions in LIFluor Enhance user's guide (PN 725824). The target RFU value was set as 62. This calibration was done whenever capillary or LIF detector or the laser was changed. After the calibration was complete and the CCF was accepted, vials of water were placed back to the "home" positions (BI:A1 and BO:A1) so that the capillary ends were placed in water to keep them wet.

**Instrument Performance Test (optional):** Instrument and gel buffer performance may be tested using dsDNA 1000 Test Mix (PN 477414) and LIFluor Enhance stain (PN 477409). Test Mix can be diluted by adding 1 mL of filtered deionized water (> 18MΩ) to the test mix vial. 5 µl LIFluor Enhance stain may be added to 6 mL of 1X rehydrated dsDNA 1000 gel buffer. Separation may be carried out following instructions in LIFluor Enhance user's guide (PN 725824).

**Preparation of Buffer Trays and Sample Trays:** Vial positions for buffer trays are indicated in Figure 3. Each "DDI Water" vials were filled with 1.5 ml double deionized (DDI) water. Waste vials were filled with 1 ml DDI water. "Gel Fill" vials were filled with 1.5 ml gel. "Gel Sep." vials were filled with 1.2 ml gel. Vial "A6" on the outlet was filled with 1 ml gel. Two sets of water dip vials were set up for minimal sample carryover and better repeatability. Injecting samples against a half-filled vial of gel buffer (BO:A6) facilitated improved precision in peak area.

Each plasmid sample was transferred to a Microvial (PN 144709) at 95 to 100 µl. Sample vials were loaded onto the sample inlet tray.

### Instrument Set up:

The "Initial Conditions" and "LIF Detector Initial Conditions" were set up as indicated in Figure 4 and Figure 5 respectively. Same set up was used for all three methods: Capillary Conditioning, Separation and Shut Down.

The time program for Capillary Conditioning was illustrated in Figure 6. Buffer vials used for conditioning were purposely positioned in row #6 to avoid moving the capillary tips over buffer vials used for sample separation.

Figures 7 and 8 show time programs for plasmid separation and shut down methods.

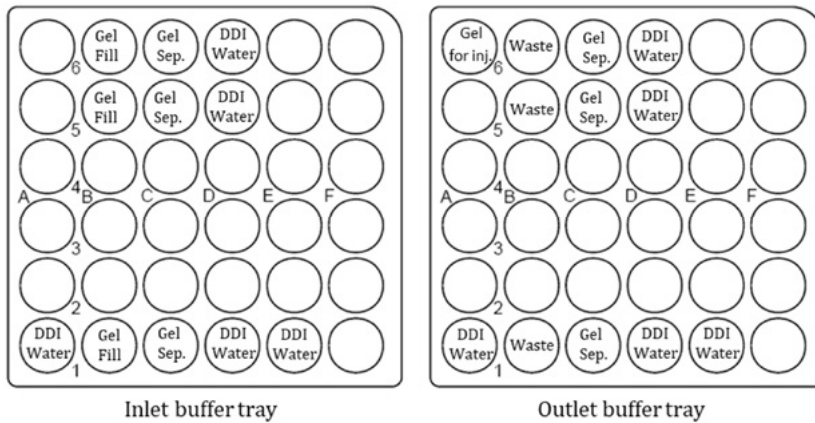


Figure 3. Schematic set up for Inlet (left) and Outlet (right) buffer trays.

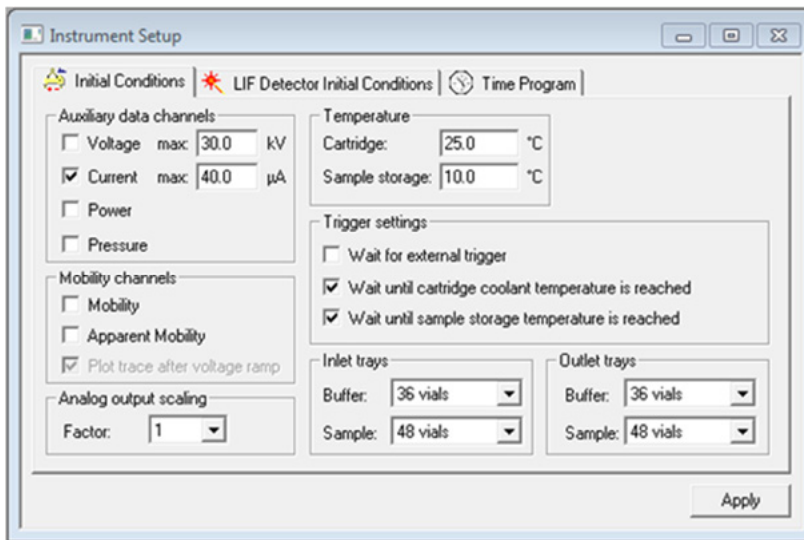


Figure 4. Settings for "Initial Conditions" tab.

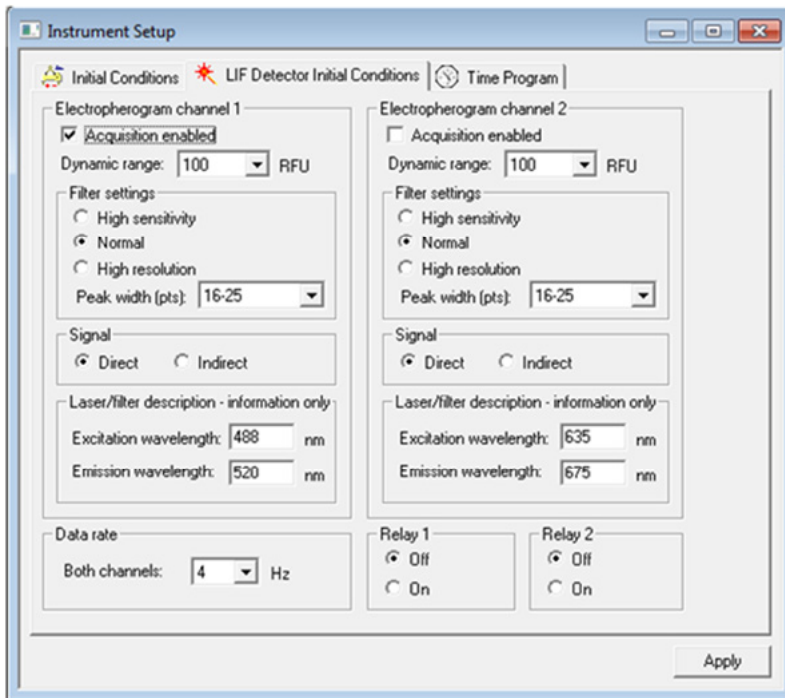


Figure 5. Settings for “LIF Detector Initial Conditions” tab.

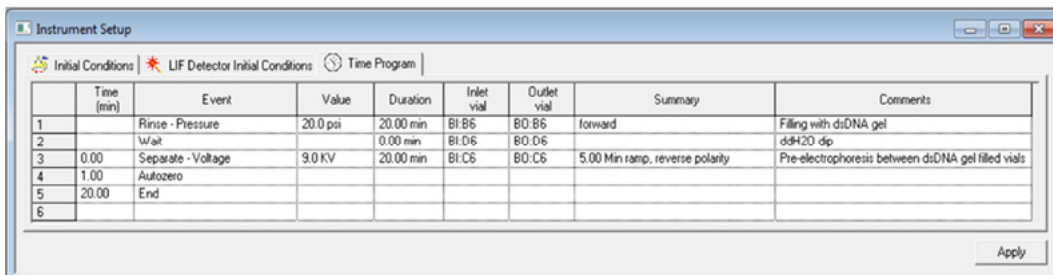


Figure 6. Time Program Settings for the Conditioning Method.

[Download this technical note](#)

Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1	Rinse - Pressure	20.0 psi	2.00 min	BI B1	BO B1	forward, In / Out vial inc 6	dsDNA Gel rinse to fill the capillary with dsDNA gel - Automatic increment every 6 runs
2	Wait		0.00 min	BI D1	BO D1	In / Out vial inc 6	water dip to clean capillary tip - Automatic increment every 6 runs
3	Inject - Pressure	0.2 psi	2.0 sec	SI A1	RI A6	Overide, forward	Sample injection with Gel in Outlet Vial
4	Wait		0.00 min	BI E1	BO E1	In / Out vial inc 6	water dip to avoid sample carry over - Automatic increment every 6 runs
5	Separate - Voltage	9.0 KV	15.00 min	BI C1	BO C1	0.17 Min ramp, reverse polarity, both, In / Out vial inc 6	Separation between dsDNA Gel - Automatic increment every 6 runs
6	Autozero						
7	15.00						
8	End						

Figure 7. Time Program Settings for the Separation Method.

Time (min)	Event	Value	Duration	Inlet vial	Outlet vial	Summary	Comments
1	Rinse - Pressure	20.0 psi	20.00 min	BI B5	BO B5	forward	Filling with dsDNA gel
2	Wait		0.00 min	BI D5	BO D5		ddH2O dip
3	Separate - Voltage	9.0 KV	10.00 min	BI C5	BO C5	5.00 Min ramp, reverse polarity	
4	Wait		0.00 min	BI A5	BO A5		
5	Laser - Off						
6	10.01						
7	End						

Figure 8. Time Program Settings for the Shut Down Method.

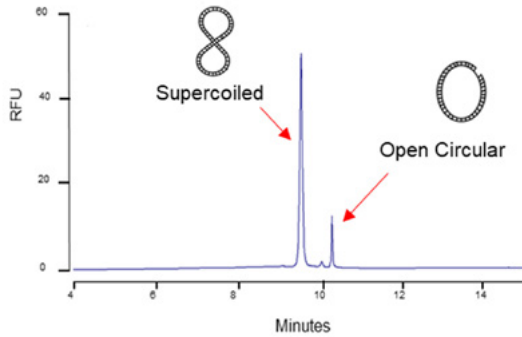
## Results

An optimized method to resolve and quantitate plasmid isoforms via capillary electrophoresis using commercially available eCAP dsDNA gel buffers is described in this technical note. When gel buffer from eCAP dsDNA 1000 kit was used without further dilution, it was found to entrap the supercoiled plasmid DNA, causing long separation times and broad circular plasmid peaks. However, by simply diluting the eCAP dsDNA 1000 gel with 1X TBE buffer, the diluted gel demonstrated the ability to provide highly resolved plasmid isoforms in a short period of time. In the case of the 5 kb test plasmid, the best results were obtained through a simple 1 to 10 dilution of dsDNA 1000 gel buffer in 1X TBE. Because the separation of the plasmid isoforms in eCAP dsDNA gels is achieved by the trapping of the supercoiled and circular isoforms of plasmid, the migration of these isoforms is slowed relative to that of the linearized molecule. This resulted in a migration pattern that is opposite to that typically observed in

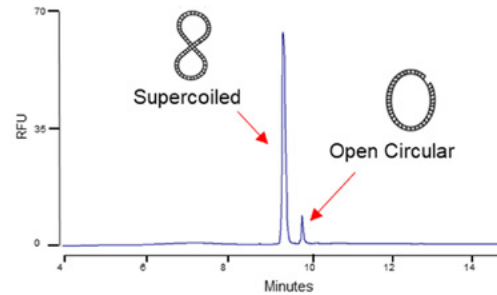
agarose gels where the super-coiled form precedes the linearized. The method presented here was optimized to work with LIF detection using LIFluor Enhance stain.

### Analysis of a 5 kb plasmid

Figure 9 shows the results obtained with a 5 kb plasmid preparation. The main isoform is the supercoiled form while the minor isoform is the open circular form. To confirm this peak assignment, a comprehensive experiment was performed where a linearized isoform was prepared by digesting the 5 kb plasmid with a single cutting restriction enzyme, BamHI-HF. In addition, a partial digestion of the plasmid with a nicking enzyme was also carried out to simulate the normal degradation of plasmid. Then, the untreated plasmid preparation, the linearized sample, the



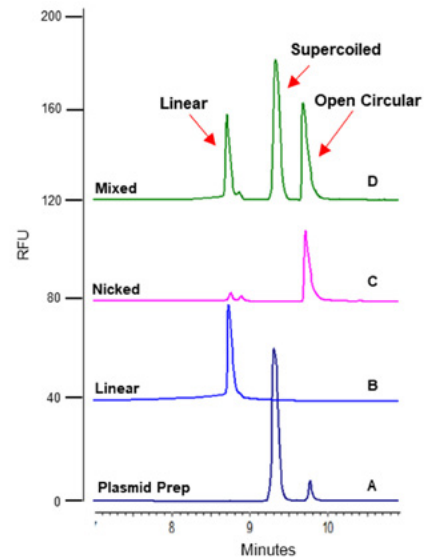
**Figure 9.** Separation of typical plasmid isoforms in a 5 kb plasmid preparation.



**Figure 10.** Separation of plasmid isoforms in a 7 kb plasmid preparation.

partially nicked sample and a mixture of the three were separated on the PA 800 Plus and compared to each other. In Figure 2, results showed that the linearized form migrated in front of the supercoiled form. The partially nicked sample contained both the supercoiled form and the open circular form as well as a trace amount of the linearized form. As expected, the amount of nicked or open circular form increased while the supercoiled form decreased in the partially nicked sample relative to those in the untreated plasmid preparation. These results indicated that this plasmid separation method is capable of monitoring plasmid degradation and also provide confirmatory information for peak assignments. In the mixed sample, all three isoforms were baseline resolved.

There is a small peak between the supercoiled form and the open circular form which could be the supercoiled dimer form. Further experiments are needed to confirm this hypothesis.



**Figure 11.** Confirmation of different isoforms in an untreated 7 kb plasmid preparation (A) using linearized (B) and nicked plasmids (C) as controls. The major peak in the untreated Plasmid Preparation (A) is the supercoiled isoform; the smaller peak is the open circular isoform.



### Analysis of a 7 kb plasmid

The plasmid analysis method using dsDNA 1000 gel, coated DNA capillary and LIFluor Enhance stain was further evaluated using a 7 kb plasmid. In Figure 10, results show a baseline resolution of the supercoiled form and the open circular form. The amount of open circular form was very low. There was no other isoform present, indicating that this is a high quality plasmid preparation. Further experiments were carried out to ensure that the small open circular peak was the nicked isoform. A linearized isoform was prepared by digesting this 7 kb plasmid with a single cutting restriction enzyme, BamHI-HF. In addition, a complete digestion of the plasmid with a nicking enzyme was also carried out. At the end of the digestion with the nicking enzyme, the reaction mixture was also heated at 80°C for 20 min to inactivate the enzyme. Then, the untreated plasmid preparation, the linearized sample, the fully nicked sample and a mixture of the three were separated on the PA 800 Plus and

compared to each other. In Figure 11, results shows the small peak in the untreated plasmid preparation was indeed the open circular form. In the fully nicked plasmid sample, there was no supercoiled isoform present. In addition, there was a small amount of linear form present possibly due to further degradation during the heat inactivation of the nicking enzyme. Once again, the results in this experiment indicates that this plasmid analysis method provides accurate assessment of the isoforms in plasmid preparation and is also capable of monitoring plasmid degradation.

### Method robustness and repeatability

The robustness and repeatability of this method was evaluated by performing 18 runs of the same untreated 5 kb plasmid sample. As shown in Figure 12, the peak profiles were very

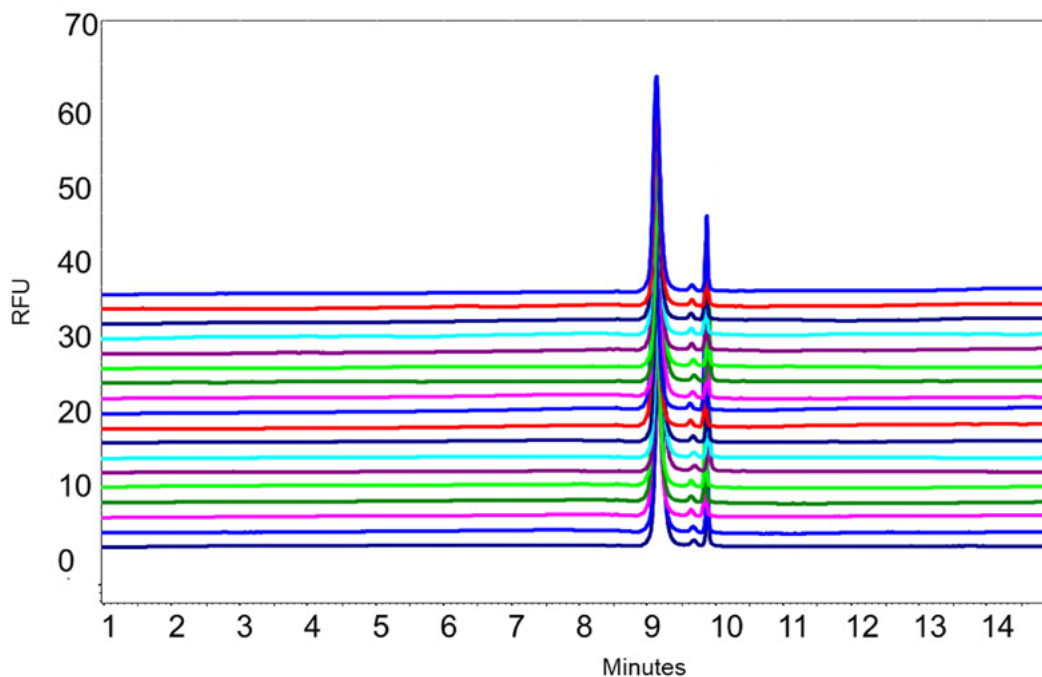


Figure 12. Overlay of 18 runs of untreated 5 kb plasmid preparation.

consistent between 18 runs. Table 1 shows the values of % RSD for migration time and % area for the supercoiled isoform obtained with the untreated 5 kb plasmid were 0.23% and 0.61% respectively.

**Table 1.** Calculation of % RSD for migration time and % area for the supercoiled isoform for 18 runs obtained with the 5 kb untreated plasmid preparation.

	Migration Time	% Area
Min:	9.11	86.28
Max:	9.18	88.54
Mean:	9.15	87.75
Std Dev:	0.02	0.54
<b>%RSD:</b>	<b>0.23</b>	<b>0.61</b>

## Discussion

This plasmid analysis method was initially optimized for the 5 kb plasmid. It also works for the 7 kb plasmid. The resolution of plasmid isoforms for the 5 kb plasmid (Figure 9) was slightly better than that for the 7 kb plasmid (Figure 10). Users may need to further optimize this method by further diluting the dsDNA 1000 gel for analysis of plasmids with sizes larger than 7 kb. Similarly, less dilution of the gel may be needed for analysis of plasmids with sizes significantly smaller than 5 kb. No matter what sizes of plasmid DNA need to be analyzed, preparing linearized and nicked plasmid isoforms as controls is a great strategy in evaluating the accuracy of the plasmid analysis method, as demonstrated in this technical note.

## Conclusions

- Fast analysis of plasmid samples with separation time within 12 minutes
- Baseline resolution of supercoiled, linear and open circular isoforms of a 5 kb and a 7 kb plasmid
- The dsDNA 1000 kit and LIFluor Enhance stain are commercially available kits for users thus streamlining the workflow
- Automated sample separation and data analysis provide rapid, reproducible and quantitative results with RSD of 0.23% for migration time and 0.61% for % area

## References

1. FDA. Points to consider on plasmid DNA vaccines for preventive infectious diseases. 1996. Docket no. 96N-0400.
2. Boardman C. and Dewald H. Plasmid purity and heterogeneity analysis by capillary electrophoresis. Poster at MSB 2010.

AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-9505 05/2019



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



## A Sensitive and Robust Plasmid Analysis Method by Capillary Electrophoresis-Laser Induced Fluorescence

Kenneth S. Cook,<sup>1</sup> Yu-ting Chen,<sup>1</sup> Sneha Chatterjee,<sup>1</sup> Jane Luo,<sup>2</sup> Tingting Li,<sup>2</sup> Marcia Santos,<sup>2</sup> Brandon Bates,<sup>2</sup> Lawrence Thompson<sup>1</sup>  
<sup>1</sup> Pfizer, Chesterfield, MO  
<sup>2</sup> SCIEX, Brea, CA

### Abstract

Plasmid DNA is an important part of the biopharma industry. It is used as a gene-delivery vehicle for DNA vaccination and as a key intermediate for processes like viral particle production for gene therapy and ex-vivo protein synthesis. Most plasmid DNA preparations contain several topological variants or isoforms including the supercoiled, open circular and linear forms of the molecule. Federal regulations require purity testing for manufactured injectable plasmid products and recommend establishing a release criterion of > 80% supercoiled content (Ref 1). Traditional agarose gel has limitations on quantitation. Capillary electrophoresis with Laser Induced Fluorescence detection (CE-LIF) provides a rapid, sensitive, reproducible and automated method for the quantitative analysis of plasmid DNA isoforms. In this poster, we describe the development of a method for plasmid analysis by CE-LIF with the use of a common dye. Results obtained with two large plasmids at 7 to 10 kb demonstrate that this method is sensitive and robust, providing baseline resolution of supercoiled, open-circular and linear plasmid isoforms within 20 minutes. It is suitable for testing plasmid purity and monitoring plasmid degradation.

### Plasmid Topology

Plasmid DNA can exist in different topological forms: covalently closed circular (CCC) or supercoiled (SC); open circular (OC) and linear (L) along with catenanes/multimers. During plasmid DNA degradation (Figure 1), the CCC form can become the OC form due to one single-stranded cleavage. When a double-stranded cleavage occurs, the plasmid DNA will become linearized.

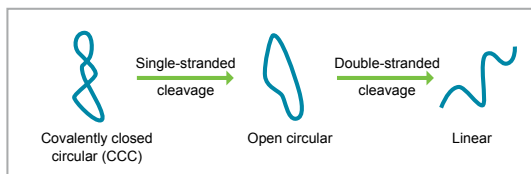


Figure 1. Plasmid Degradation.

### Agarose Gel

Agarose gel is the standard method for analysis of plasmid topology. In Figure 2, 200 ng of a plasmid sample (Lane "P2") was run on a 1% TAE agarose gel with ethidium bromide (PN 161-0433, BioRad). About 835 ng of 1 kb plus DNA ladder (PN SM1331, Thermo) was loaded in Lane "M". The supercoiled (SC) and open circle (OC) generated good, sharp bands. The linear (L) and multimer forms were barely detectable. Therefore, although the agarose gel method is easy to do, it has a high detection limit. A capillary electrophoretic method would improve resolution and detection, providing better quantitative analysis.

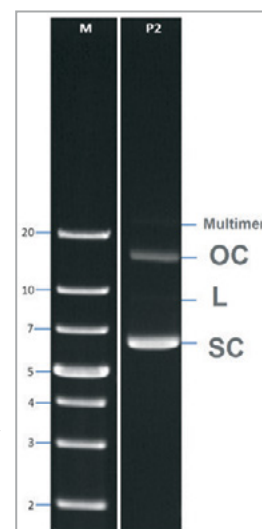


Figure 2. Plasmid analysis by agarose gel.

### Capillary Electrophoresis

Capillary electrophoresis analysis for plasmid DNA was performed on a PA800 Plus Pharmaceutical Analysis System (SCIEX) equipped with an LIF detector. A dsDNA 1000 kit (PN 477410) consists of a sieving matrix in tris-borate-EDTA, pH 8.3, buffer (TBE), a coated capillary and Orange G marker was from SCIEX.

### Vial Trays and Separation Method

Vial positions and separation timed program for the optimized plasmid analysis method are shown in Table 1 and Table 2. Cartridge temperature was set to 20° C. Sample storage temperature was set to 15° C.

[Download this technical note](#)

Water		Gel Separation	Gel Rinse		
Water		Gel Separation	Gel Rinse		
Water		Gel Separation	Gel Rinse		
Water		Gel Separation	Gel Rinse		
Water		Gel Separation	Gel Rinse		
Water	Gel Rinse	Gel Separation	Methanol Rinse		

BI (Inlet Buffer Tray)

Water	Gel	Gel Separation	Water Waste		
Water	Gel	Gel Separation	Water Waste		
Water	Gel	Gel Separation	Water Waste		
Water	Gel	Gel Separation	Water Waste		
Water	Gel	Gel Separation	Water Waste		
Water	Gel	Gel Separation	Water Waste		
Water	Water Waste	Gel Separation	Water Waste		

BO (Outlet Buffer Tray)

**Table 1.** Vial positions for optimized plasmid method.

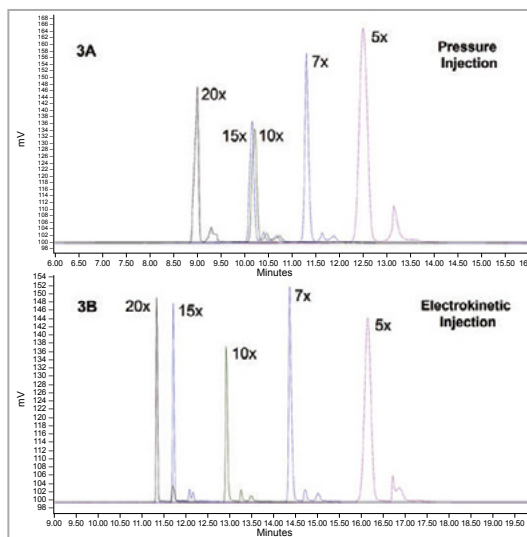
Time (min)	Event	Value	Duration	Inlet	Outlet	Summary
	Rinse - Pressure	10.0 psi	20.00 min	BI:D2	BO:D2	Forward, In/Out vial inc 6
	Wait		0.00 min	BI:A2	BO:A2	In/Out vial inc 6
	Inject Pressure	2.0 KV	10.0 sec	SI:A1	BO:B2	Override, forward
	Wait		0.00 min	BI:A2	BO:A2	In/Out vial inc 6
0.00	Separation - Voltage	7.8 KV	20.00 min	BI:C2	BO:C2	0.17 min ramp, reverse polarity, In/Out vial inc 6
1.00	Autozero					
20.00	End					

**Table 2.** Separation timed program for optimized plasmid method.

## Optimization of Sample Injection and Separation Matrix Conditions

Although different plasmid DNA topoisomers have similar molecular weights, their tertiary structures are different, allowing the gel buffer to sieve. Sieving was achieved by diluting the gel buffer to obtain the best resolution. During initial experiments, the gel buffer was diluted with 1xTBE at dilution factors of 5x, 7x, 10x, 15x and 20x. It was determined that gels diluted at 7x and 10x generated the best results (Figure 3A and 3B). The 10x dilution gel was used for further optimization experiments from this point on.

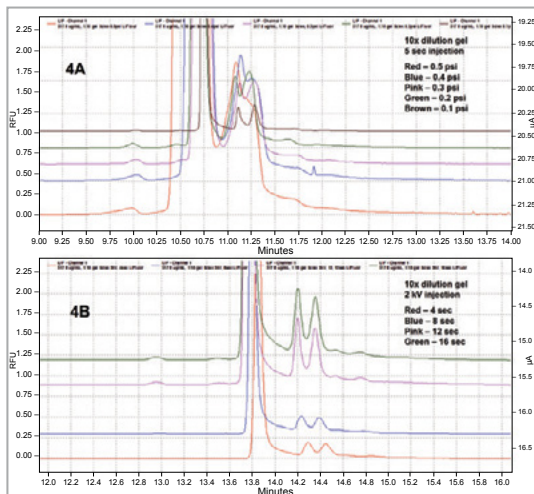
Since pressure injection may load more sample matrix than electrokinetic injection, an evaluation of pressure (Figure 4A) and electrokinetic injection (Figure 4B) was done with 10x dilution gel, LIFluor Enhance stain and 8 ng/ $\mu$ L plasmid DNA sample. Results in Figure 4A showed that better resolution was obtained with lower injection pressure with best resolution at 0.1 psi for 5 seconds. Much better resolution was achieved when electrokinetic injection was used to introduce the samples (Figure 4B). Among the conditions used with electrokinetic injection, 12 seconds injection time at 2.0 kV produced the best results.



**Figure 3.** Varying gel dilution with pressure (0.2 psi for 4 sec; Panel 3A) or electrokinetic (2 kV for 5 sec; Panel 3B) injections. LIFluor Enhance stain (PN 477409, SCIEX) was used.

[Download this technical note](#)

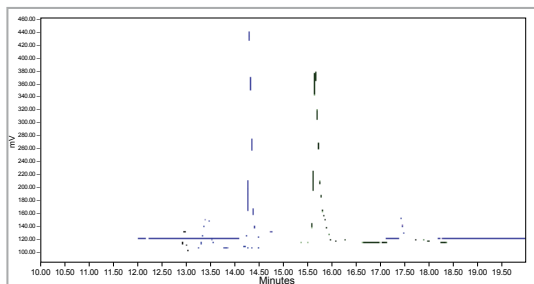




**Figure 4.** Effect of varying injection pressure (Panel 4A) and varying electrokinetic injection time (Panel 4B). LIFluor Enhance CE stain was used.

## Different Fluorescent Dyes

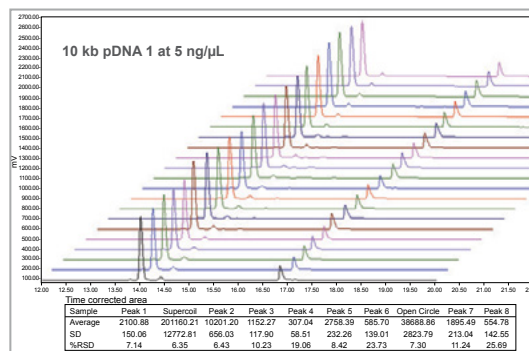
LIFluor Enhance, Ethidium Bromide (EtBr), Sybr-gold (PN S11494, Thermo) and YOYO-1 were evaluated. Sybr-gold generated the best separation of supercoiled and open circle forms (Figure 5).



**Figure 5.** When LIFluor Enhance (black) and EtBr (blue) were used, sample concentration was at 8 ng/μL; LIF dynamic range was 100 RFU; injection was 2 kV for 5 sec. For Sybr-gold (purple), sample was 5 ng/μL; detection range was 1000 RFU; injection was 2 kV for 10 sec. For YOYO-1 (green), sample was 8 ng/μL; detection range was 1000 RFU; injection was 2 kV for 5 sec.

## Robustness

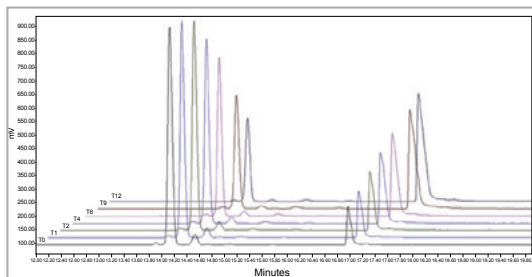
Further optimization involved adding methanol rinse as the first step in conditioning. Methanol rinse improved separation profiles, extended the capillary life. Methanol rinse between injections or at specific intervals did not enhance resolution (data not shown). Robustness of optimized method was tested by running the same sample multiple times. Results in Figure 6 demonstrates consistent peak profiles for 21 consecutive runs with different topological isoforms baseline resolved. Percent RSD values of time corrected area for the supercoiled and the open circle peaks were below 8%.



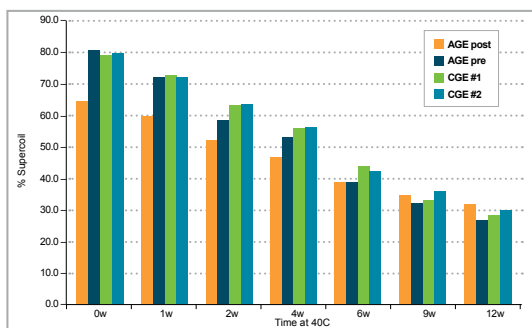
**Figure 6.** Twenty-one consecutive runs of a 10 kb plasmid using optimized method. Sybr-gold stain was used.

## CE for Plasmid Stability Monitoring

pDNA 1 (10 kb) and pDNA 2 (7 kb) were stressed at 40° C to produce different isoforms. Samples were then analyzed by different topology methods: capillary gel electrophoresis (CGE) and agarose gel electrophoresis (AGE). Results in Figure 7 and 8 indicate CGE can show topology changes. The supercoiled form decreased while open circle form increased over stress treatment (Figure 7). Results from CGE were consistent with those from AGE. The topoisomorph percentages were not affected by the methanol rinse in CGE (Figure 8).



**Figure 7.** Robustness in monitoring degradation of pDNA 2 (7 kb) samples stressed at 40°C from zero (T0) to 12 (T12) weeks.



**Figure 8.** Robustness in monitoring degradation of pDNA 1 (10 kb) samples stressed at 40°C from zero (T0) to 12 (T12) weeks. CGE #1 is no methanol rinse and CGE #2 includes methanol rinse. "AGE post" was stained after running while "AGE pre" contained stain in the agarose gel.

## Conclusions

- A sensitive and robust CE-LIF method for plasmid analysis was developed.
- Optimizing the gel dilution allowed plasmid isoforms to be baseline resolved and quantitated.
- Electrokinetic injection performed better than pressure injection.
- Intercatation of different dyes have different impact on the topoisomere separation. Sybr-gold produced the best separation profile.
- Addition of a methanol rinse as the first step in conditioning extended number of injections per sequence and increased the capillary life time to over 100 injections. The topoisomere percents were not affected by the methanol rinse.
- Experiments with stressed plasmid samples showed CE-LIF can quantitate plasmid topoisomers.

## References

1. FDA. Points to consider on plasmid DNA vaccines for preventive infectious diseases. 1996. Docket no. 96N-0400.



# Selection of aptamers

## CE-SELEX: Rapid Aptamer Selection Using Capillary Electrophoresis.

Sean K. Dembowski, Michael T. Bowser  
University of Minnesota, Minneapolis, Minnesota

### Introduction

Systematic evolution of ligands by exponential enrichment (SELEX) is a process for obtaining aptamers (short DNA or RNA sequences) with high affinity and selectivity for a molecular target. While aptamer applications are growing exponentially, the time and difficulty of the current SELEX process used to isolate aptamers are limiting widespread adoption.

Recently, the high resolving power of capillary electrophoresis (CE) has been used to select high-affinity aptamers, reducing the process from weeks to only a few days, using a process known as CE-SELEX. In CE selections, nucleic acid sequences that bind the target undergo a mobility shift, migrating at a different rate allowing separation and collection of high-affinity aptamers from the non-binding sequences.

CE-SELEX with the SCIEX P/ACE™ MDQ Plus offers a number of advantages over conventional selection protocols due to the dramatic enrichment rate and selection stringency.

- CE-SELEX can isolate high-affinity aptamers in fewer rounds and without tedious negative selection compared to conventional SELEX methods, **shortening a several-week process down to as little as a few days.**<sup>1</sup>
- CE selection occurs in free solution, eliminating the need for filtration or solid-phase attachment of the target. This **increases the number and types of viable targets** (including targets smaller than the aptamer itself).
- Many aptamers selected using CE-SELEX exhibit similar or **better affinity for their targets** than aptamers selected using conventional methods.<sup>2</sup>
- CE-SELEX offers **flexibility to manipulate the selection stringency** by varying target concentration, separation parameters, and collection window timing.
- Contrary to conventional selections, sequence motifs are rare in aptamers selected using CE-SELEX, allowing **more flexibility in developing an optimized aptamer sequence.**
- CE-SELEX is **compatible with many non-natural nucleic acid libraries** and modifications that cause issues for other SELEX techniques.

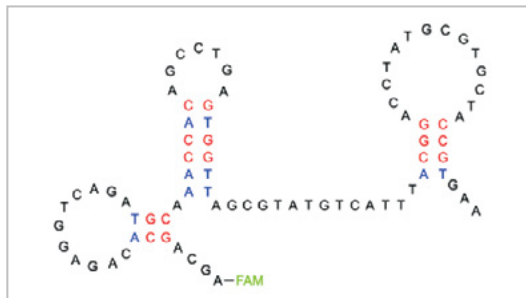


Figure 1: Secondary structure of a DNA aptamer selected using CE-SELEX to have affinity for neuropeptide Y. (Mendonça, S. D., Bowser, M. T., *J. Am. Chem. Soc.*, 2005, 127, 9382–9383.)

- CE-SELEX **easily supports limited samples**. Target concentrations as low as 1 pM have successfully been used,<sup>3</sup> and novel nanoVials allow sample collection in as little as 5 µl buffer volume.

### Overview

The word aptamer comes from the Latin root “aptus” meaning “to fit.” This is appropriate since aptamers are single-stranded deoxyribonucleic acid (ssDNA) or ribonucleic acid (RNA) sequences that fold into unique structures, allowing them to bind target molecules with high affinity and selectivity. Depending on their sequence, single-stranded nucleic acids can fold into a variety of loops, hairpins, and bulges to generate a wide range of structures (Fig. 1). The plethora of structures available makes nucleic acids an attractive combinatorial library since sequences can be isolated with affinity for virtually any molecular target.

In the SELEX method, a randomized nucleic acid pool is incubated with a target. Sequences with affinity for the target are separated, amplified, and incubated again in an iterative process. The step of separating binding from non-binding sequences is often carried out by affinity chromatography or filtration-based methods. Unfortunately, both of these methods are time-consuming and inefficient. Isolating aptamers using capillary electrophoresis (CE) combines extreme resolving power and

p1

[Download this technical note](#)

highly stringent yet flexible selection conditions allowing high-affinity aptamers to be obtained in only 2-4 rounds of selection.<sup>4,5</sup> Compared to conventional methods, which require 8-15 rounds and tedious negative selection, CE is a simple and efficient technique that dramatically decreases the time and resources required for isolation of new aptamer sequences.

The CE-SELEX process is illustrated in Figure 2. The target molecule is incubated with a random sequence nucleic acid library. Several nanoliters of this incubation mixture are injected onto a capillary and separated using free zone CE. Nonbinding oligonucleotides migrate through the capillary with the same mobility, regardless of their length or sequence. Complexing the target changes the size and charge of binding sequences, causing them to migrate as a separate fraction. This fraction of binding sequences is collected at the outlet of the capillary. As in conventional SELEX selections, the fraction containing binding sequences is amplified, purified, and made single stranded for further rounds of enrichment. The process is repeated until no further improvement in affinity is observed between rounds.

## Materials

All solutions were prepared in 0.2 µm-filtered water, unless nuclease-free water (Invitrogen, Carlsbad, CA) is specified. All DNA was obtained from IDT DNA (Coralville, IA). All other equipment and reagents were obtained from Thermo Fisher (Waltham, MA) in analytical grade unless otherwise specified.

### Capillary Electrophoresis Separation

1. An automated P/ACE™ MDQ Capillary Electrophoresis instrument (Beckman Coulter, Fullerton, CA) equipped with all standard features plus a laser module to facilitate LIF detection. The procedure is directly transferable to the automated P/ACE™ MDQ Plus equipped with a solid state laser module (SCIEX, Redwood Shores, CA).
2. Bare fused silica eCap Capillary with 50 µm i.d. and 375 µm o.d. (SCIEX, Redwood Shores, CA).
3. Nuclease-free centrifuge tubes, 0.6- and 1.5 mL sizes (Dot Scientific, Burton, MI).
4. Sterile, polyethylene-coated filter pipette tips to avoid contamination (VWR, Radnor, PA).
5. 5' FAM labeled ssDNA library consisting of 40 random bases flanked by 20-base PCR primer regions. Dilute to 400 µM in nuclease-free water and store in the dark at -80° C.

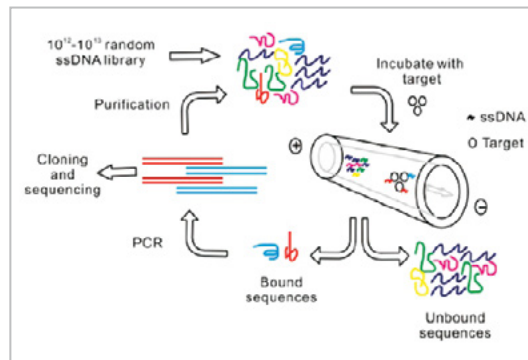


Figure 2: Schematic representation of the CE-SELEX process. A random sequence ssDNA library is incubated with the target molecule, then binding sequences are separated using CE, PCR amplified, and purified, generating a new pool suitable for further cycles of enrichment. High-affinity aptamers are typically obtained after 2-4 rounds of selection. (Yang, J.; Bowser, M. T., *Anal. Chem.*, 2013, 85 (3), 1525-1530.)

6. 5x TGK separation buffer consisting of 125 mM Tris-HCl, 960 mM glycine (Sigma-Aldrich, Saint Louis, MO), and 25 mM KH<sub>2</sub>PO<sub>4</sub> (Avantor, Center Valley, PA), pH 8.3. Store at 2-8° C. Dilute to 1x working concentration as necessary.
7. Sample buffer for incubation. Sample buffer should match the anticipated application conditions (e.g., pH, ionic strength, ionic composition, etc.) as closely as possible to maximize aptamer performance. Prepare buffer stock of 5x concentration and dilute to working concentration as needed. Store at 2-8° C.

*Note: Working sample buffer should contain at least 5 mM K<sup>+</sup> to allow formation of DNA G-quadruplex motif. If conditions for anticipated aptamer usage are unknown, use 1x TGK as sample buffer.*

### Polymerase Chain Reaction (PCR)

1. Thin walled 0.5 mL PCR tubes (Eppendorf, Westbury, NY).
2. Automated thermal cyclers (Eppendorf, Westbury, NY).
3. 5' FAM labeled forward DNA primer. Dilute to 60 µM in nuclease-free water and store in the dark at -80° C.
4. 5' biotin labeled reverse DNA primer. Dilute to 60 µM in nuclease-free water and store at -80° C.
5. Set of 100 mM deoxynucleotide triphosphates (dNTP's) (Invitrogen, Carlsbad, CA). Store at -20° C.
6. 10x Thermo pol buffer and 5000 U/mL Taq polymerase (New England Biolabs, Ipswich, MA). Store at -20° C.



7. Solution of 25 mM MgCl<sub>2</sub> (Sigma-Aldrich, Saint Louis, MO) prepared in nuclease-free water. Store at 2 - 8° C.

### Agarose Gels

1. Gel box for horizontal 14 cm x 11 cm agarose gel with power supply and comb (at least 12 sample wells).
2. Ultraviolet Transilluminator for gel imaging (Sigma-Aldrich, Saint Louis, MO).
3. Molecular biology grade agarose powder.
4. 10× TBE buffer prepared with 1.0 M Tris - HCl, 0.9 M boric acid, and 20 mM disodium EDTA (Sigma-Aldrich, Saint Louis, MO) at pH 8.3. Dilute to 0.5x as needed. Store in the dark at room temperature.
5. Ethidium bromide (Sigma-Aldrich, Saint Louis, MO), diluted to 10 mg/mL. *Ethidium bromide is a known mutagen. Gloves, lab coat, and goggles should be worn when handling.* Store in the dark at room temperature.
6. 6× Blue-orange dye (New England Biolabs, Ipswich, MA). Store at -20° C.
7. 25 base pair DNA molecular weight ladder (Invitrogen, Carlsbad, CA). Store in the dark at -20° C.

### Streptavidin Column Purification

1. Amicon 0.5 mL centrifugal filter, 10 kDa cutoff (EMD Millipore, Billerica, MA).
2. Centrifuge, capable of 14,000 x g.
3. Streptavidin agarose resin. Store at 2-8° C.
4. Streptavidin-binding buffer prepared with 50 mM NaCl, 10 mM Tris-HCl, 1 mM disodium EDTA (Sigma-Aldrich, Saint Louis, MO) in nuclease-free water. Store at 2-8° C.
5. 10 mL Poly-prep chromatography columns with stopper and lid (Bio-Rad, Hercules, CA).
6. Solution of 0.15 M NaOH. Store at 2-8° C.
7. Solution of 0.15 M acetic acid. Store at 2-8° C.

### Sequencing

1. Adjustable temperature water bath and shaking incubator.
2. Unlabeled forward and reverse primers. Dilute each to 60 μM in nuclease-free water and store at -80° C.
3. TOPO TA cloning kit with One Shot TOP10 chemically competent *E. coli* (Invitrogen, Carlsbad, CA).
4. Standard LB selection plates containing 50 μg/mL ampicillin or kanamycin (Sigma-Aldrich, Saint Louis, MO).

5. PureLink Plasmid Miniprep Kit (Invitrogen, Carlsbad, CA).

## Methods

### Identification of the Collection Window

1. The migration of the oligonucleotide library and target should be determined separately on CE. Common initial separation conditions are as follows: 1x TKG separation buffer, 100 nM ssDNA or 5 μM target in 1x sample buffer, 40 cm capillary (30 cm to detector), 30 kV normal polarity (cathode at outlet) separation, 4 s injection at 1 psi, LIF detection.

*Note: If target is not fluorescent, target migration should be monitored by UV/Vis absorbance detection rather than LIF detection.*

2. Electropherograms in which the target peak exhibits significant tailing or a peak for the target is not observed may suggest significant adsorption of the target to the capillary surface. In such cases, a coated capillary should be used in place of a bare fused silica capillary to reduce these surface interactions. Coated capillaries with pre burned windows are available through SCIEX ([www.SCIEX.com](http://www.SCIEX.com)). Use reversed polarity (anode at outlet) for coated capillaries.
3. Modify the separation conditions to achieve an adequate separation between the unbound library peak and the target peak by changing separation voltage, capillary length and/or inner diameter, buffer components, and injection volume.

*Note: Very high ionic strength in the sample buffer relative to separation buffer may cause DNA peak shape to deteriorate due to destacking. If poor DNA peak shape makes separation difficult or variable, decrease sample buffer ionic strength.*

4. **For an uncoated capillary:** Separate DNA library in normal polarity using the conditions optimized previously. Note the time of the leading edge of the unbound sequences, then subtract 10-20 seconds. This is  $t_{det}$ , the time at which the end of the collection window passes the detector (Fig. 3). (If no DNA is successfully collected, substitute the DNA leading edge time for  $t_{det}$  without subtracting 10-20 seconds.)

**For a coated capillary:** Separate DNA library in reversed polarity using the conditions optimized previously. Note the time of the end of the unbound sequence peak, then add 10-20 seconds. This is  $t_{det}$ , when the start of the collection window passes the detector (Fig. 4). (If no DNA is successfully collected, substitute the end time of the DNA peak for  $t_{det}$  without adding 10-20 seconds.)

Note: In the rare case that the net positive charge of the target exceeds the net negative charge of the DNA sequence and a coated capillary is used, the net positive complex should migrate toward the cathode. In this case, use a coated capillary but follow the directions for fraction collection below as if using an uncoated capillary and use a  $t_{det}$  of 20 min if no unbound DNA peak is observed.

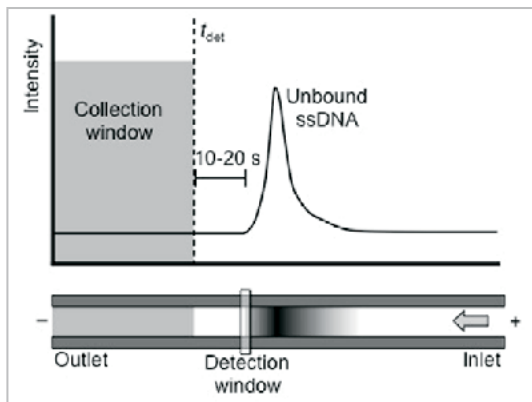


Figure 3: CE-SELEX collection strategy when using an uncoated capillary. All species migrate toward the cathode, with DNA–target complexes generally migrating off the capillary before unbound DNA. Collection starts immediately and ends at  $t_{det}$ , 10–20 seconds before the unbound DNA peak. (Adapted from Mosing, R. K.; Bowser, M. T., CE-SELEX: Isolating Aptamers Using Capillary Electrophoresis. In *Nucleic Acid and Peptide Aptamers*. Mayer, G., Ed. Humana Press: New York, NY, 2009; p 33.)

5. The above times determine collection window in relation to the detection point. To determine when the edge of the collection window reaches capillary outlet, use the equation:

$$t_{out} = \frac{(L_T)(t_{det})}{L_D} \quad (1)$$

where  $t_{out}$  is the time that the end of the collection window reaches the outlet,  $L_T$  is the total length of the capillary, and  $L_D$  is the length from inlet to detector. (For reversed polarity separation on a coated capillary,  $t_{out}$  is instead when the start of the collection window reaches the outlet.)

### Fraction Collection using Capillary Electrophoresis

1. Heat an aliquot of ssDNA library (or purified ssDNA sample from previous round) to 72° C for 5 min and allow to cool to room temperature to ensure reproducible DNA folding. Bring an aliquot of target stock to room temperature. A high target stock concentration should be used to minimize DNA dilution when preparing samples in subsequent rounds.

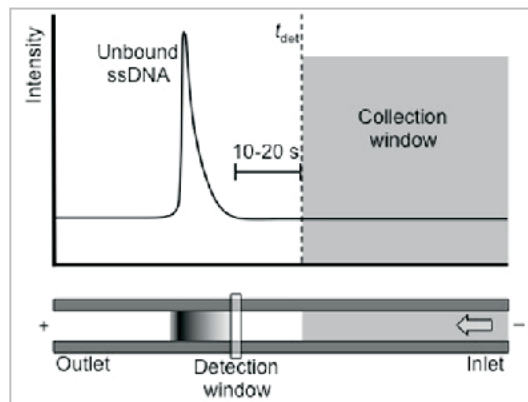


Figure 4: CE-SELEX collection strategy when using a coated capillary. DNA migrates toward the anode, with DNA–target complexes generally migrating off the capillary after unbound DNA. Samples are separated until  $t_{det}$ , 10–20 seconds after the end of the unbound DNA peak, then complexes are collected by rinsing. (Adapted from Mosing, R. K.; Bowser, M. T., CE-SELEX: Isolating Aptamers Using Capillary Electrophoresis. In *Nucleic Acid and Peptide Aptamers*. Mayer, G., Ed. Humana Press: New York, NY, 2009; p 33.)

2. Combine and dilute ssDNA, 5x sample buffer, and target stock accordingly to obtain a 20  $\mu$ L sample of 1x sample buffer, 0.1–10 nM target, and as concentrated ssDNA as possible. Typically 200–300  $\mu$ M ssDNA library is used in the first selection round, while limited PCR yield allows samples up to 1–5  $\mu$ M purified ssDNA in subsequent rounds. Incubate in the dark for 20 min to allow binding to occur. Incubation temperature should match temperature of the anticipated aptamer application, if known.

Note: Higher concentrations ssDNA vs. the target ensures competition for binding sites. We advise starting with 1 nM of target and only increasing if DNA is not present in the collection fraction.

3. **For an uncoated capillary:** Inject the sample mixture onto the capillary for 4 s at 1 psi. Separate using the normal polarity parameters optimized previously. Perform the separation directly into an outlet vial containing 60  $\mu$ L of separation buffer, and stop the separation at  $t_{out}$ . The outlet vial now contains the bound sequences of interest. Rinse the capillary to waste with buffer for 5 min at 20 psi to remove all unbound sequences. To sample a greater number of sequences in the first selection round, repeat this step 2–4 additional times using the same outlet vial.

**For a coated capillary:** Inject the sample mixture onto the capillary for 4 s at 1 psi. Separate using the reversed polarity parameters optimized previously. Stop the separation at  $t_{out}$ .

The outlet vial now contains the unbound sequences. Dip the outlet into nuclease-free water (programmed wait, 30 s) to remove any DNA adhered to the outside of the capillary, then rinse buffer through the capillary (containing the bound fraction) into a clean collection vial containing 50  $\mu$ L of buffer with a 20 psi rinse for 3 minutes. To sample a greater number of sequences in the first selection round, repeat this step 2-4 additional times using the same collection vial.

### Polymerase Chain Reaction (PCR)

- Using filter pipette tips, prepare a PCR master mix containing the following:

498  $\mu$ L nuclease-free water  
 300  $\mu$ L 25 mM  $MgCl_2$   
 100  $\mu$ L 10 $\times$  thermo pol buffer  
 2.5  $\mu$ L each 100 mM dNTP (dATP, dCTP, dGTP, dTTP)  
 8.5  $\mu$ L 60  $\mu$ M forward primer with 5' 6-FAM label  
 8.5  $\mu$ L 60  $\mu$ M reverse primer with 5' biotin label

*Note: Master mix is particularly susceptible to contamination. Prepare master mix fresh each day and take extra precaution to avoid contaminating stocks or master mix.*

- Distribute 92.5  $\mu$ L of the PCR master mix into nine thin walled PCR tubes using a filter pipette tip.
- Add 7  $\mu$ L nuclease-free water to one tube with a filter pipette tip. Label this tube as the negative control and set aside.
- To each remaining tube, add 7  $\mu$ L of the collected CE fraction using a filter pipette tip.
- Place the PCR tubes in a thermal cycler. Heat the samples to 94 $^\circ$  C for 2 min. Keeping all other tubes at 94 $^\circ$  C, add 0.5  $\mu$ L of 5,000 u/mL Taq polymerase to each sample one at a time using a filter pipette tip, starting with the control.
- Run the desired PCR method according to standard protocol. A typical method is 25 cycles of 94 $^\circ$  C denaturation (30 s), 53 $^\circ$  C annealing (30 s), 72 $^\circ$  C extension (20 s, 5 min for final cycle). However, adjustment may be needed based on primer length and melting temperature, collected DNA concentration, master mix component concentrations, and specific thermal cycler used.

### Agarose Gels

- Prepare a 1.5% agarose gel by adding 1.5 g agarose to 100 mL 0.5 $\times$  TBE buffer in an Erlenmeyer flask. Swirl to mix, then heat the mixture in microwave for ~2 min or until the agarose is in solution, making sure that solution does not boil over. Remove solution from microwave and allow to cool until container can be handled comfortably (~60 $^\circ$  C). Add 1  $\mu$ L of 10 mg/mL ethidium bromide. *Do not add the ethidium bromide until the solution has cooled to 60 $^\circ$  C or less to prevent harmful aerosols.* Swirl the solution to mix and pour into UV

transparent gel tray. Immediately place comb in gel. Allow the gel to set for 30 min at room temperature or 10 min in the refrigerator.

- Prepare the samples to be run on the agarose gel. First, add 1  $\mu$ L 6 $\times$  blue-orange dye to 11 microcentrifuge tubes. Add 5  $\mu$ L 25 base pair DNA ladder to two of these tubes and set them aside. To the remaining nine tubes, add 5  $\mu$ L from each PCR sample and control. Mix by pipetting.

*Note: DNA ladder may require additional dilution before use. See specific ladder instructions for recommended concentration.*

- Carefully remove the comb from the gel. Submerge the gel (still in the UV transparent tray) into the gel box filled with 0.5 $\times$  TBE buffer and use a pipette tip to remove bubbles from wells.
- Load 5  $\mu$ L of each sample into its own well. The two ladder samples should be loaded near opposite sides of the gel to correct any lateral field differences.
- Complete the gel apparatus assembly and plug into the power supply. Run at 200 V for 40 min or until a good separation between the yellow, purple, and blue dyes is achieved. Do not allow any of the dyes to migrate off the gel. The voltage connections should be confirmed by observing bubble formations at the electrodes. DNA will migrate toward the positive electrode.
- Once the gel run is complete, disconnect from the power supply. Carefully place the gel on the UV box for imaging. *Do not look directly into the UV box without UV eye protection.* The gel should have a single sharp band at the expected DNA length. Additional bands in the sample lane suggest contamination, while fuzzy or smeared bands may require reduction of initial DNA or PCR cycles used.

*Note: Negative controls should be scrupulously carried out every round to confirm that the DNA obtained is the desired product. Additionally, a positive PCR control using 50-500 pM ssDNA library may be run if no sample bands are observed or if contamination of reagents is suspected.*

### Streptavidin Column Purification

- Vortex the streptavidin agarose resin stock to evenly distribute the settled beads into solution. Place 300  $\mu$ L streptavidin agarose resin and 300  $\mu$ L streptavidin binding buffer into a stoppered poly prep chromatography column. Vortex briefly then drain liquid to waste.
- Restopper and add 300  $\mu$ L streptavidin binding buffer to all PCR samples except control tubes. Cap and incubate for 20 min at room temperature, vortexing periodically.
- Remove stopper and rinse column with 500  $\mu$ L streptavidin binding buffer about ten times to remove excess PCR reagents.

*Note: Normal flow through the column may be very slow. To speed up flow, press a gloved hand or finger down on top of the column to apply pressure. Do not to disturb the plug of beads when pressure is released.*

4. Restopper and add 200  $\mu\text{L}$  0.15 M NaOH. Vortex thoroughly and incubate at 37° C for 20 min to break the hydrogen bond network between the double stranded DNA without disrupting the biotin–streptavidin complex.
5. Elute the liquid (containing the ssDNA sequences) into a clean 1.5 mL centrifuge tube. Immediately add 200  $\mu\text{L}$  0.15 M acetic acid to the centrifuge tube to neutralize the solution.
6. Repeat **Steps 4-5** one more time, eluting ssDNA sequences and adding acetic acid into the same tube.
7. Split the ssDNA (~800  $\mu\text{L}$ ) into two Amicon centrifugal filters. Centrifuge for 15 min at 14,000 x g. Add 400  $\mu\text{L}$  nuclease-free water to each filter and centrifuge again.
8. Invert each filter into a clean tube (provided with filters) and centrifuge for 7 min at 1,000 x g to elute purified ssDNA sample. Small sample volume (~50  $\mu\text{L}$ ) should be visible. Store in the dark at -80° C. The ssDNA sample is now ready for the next round of selection.

*Note: For long-term storage of ssDNA, samples should be dried prior to freezing via vacuum centrifugation (50° C, 3 hrs) or lyophilization. Typically one tube is dried for future dissociation constant measurements and/or sequencing while the other is used for the next round of selection.*

*Note: Prior to dissociation constant measurement and future rounds of selection, the approximate concentration of purified PCR product should be determined. This can be done by looking at peak height on CE, using ssDNA library as an external standard.*

### Bulk Dissociation Constant Measurements

1. Bulk dissociation constants can be determined using affinity capillary electrophoresis (ACE). Dilute purified ssDNA pool to 40 nM in 1 $\times$  sample buffer. Heat DNA to 72° C for 5 min and allow it to cool to room temperature to ensure reproducible room temperature conformations.

*Note: Concentration of DNA usually does not need to be known exactly; however, concentration must be high enough to produce consistent signal intensity in CE runs but should ideally be much smaller than  $K_d$ . LIF detection must be used for these measurements, as UV absorbance detection is not sensitive enough for these small DNA concentrations.*

2. Prepare serial dilutions of target in sample buffer, typically eight or more concentrations ranging from 50 nM to 10  $\mu\text{M}$ . For improved fitting, the highest concentration should be at least 10x greater than expected dissociation constant. Add 10  $\mu\text{L}$  of ssDNA pool to 10  $\mu\text{L}$  of each diluted target sample in a separate 0.6 mL tube. Also prepare two blank samples by adding 10  $\mu\text{L}$  ssDNA pool to 10  $\mu\text{L}$  of sample buffer. Mix samples and allow to incubate for 20 min (at temperature used in selection incubation) to allow binding to occur.

*Note: Samples should contain equal concentrations of all species except the target molecule. If the target stock contains additional compounds not present in the sample buffer (such as a surfactant or buffer additives), then these compounds should be equalized across all samples or removed from the target stock prior to sample preparation.*

3. Analyze each sample on CE using identical conditions to selection. Monitor the height of the free aptamer peak for each sample. The free aptamer peak height should be greatest in the blank samples and decrease as the target concentration increases (due to an equilibrium shift from the unbound aptamer toward the aptamer–target complex).
4. Assuming the aptamer concentration is much lower than the target concentration, the dissociation constant ( $K_d$ ) can be estimated by fitting the heights of the unbound peak to the following equation:

$$\frac{I_0 - I}{I_0} = \frac{[\text{target}]}{K_d + [\text{target}]} \quad (2)$$

where  $I_0$  is the height of the unbound aptamer peak in the absence of target (blank),  $I$  is the height of the unbound aptamer peak of a particular sample,  $[\text{target}]$  is the concentration of the target in that sample, and  $(I_0 - I)/I_0$  is the fraction of DNA bound.

*Note: Eqn. 2 is only accurate so long as aptamer concentration is at least an order of magnitude smaller than  $K_d$ . If this is not the case, a more complex equation<sup>6</sup> can be used to fit the data and determine  $K_d$ .*

### Cloning and Sequencing

*Note: Before beginning, ensure that all directions and requirements pertaining to TOPO TA cloning kit, One Shot E. coli cells, PureLink plasmid miniprep kit, and DNA sequencing standard protocol have been read and understood thoroughly.*

1. Amplify the pool to be sequenced (usually the round(s) with the lowest bulk dissociation constant) by following **Polymerase Chain Reaction (PCR)** directions above, except unlabeled forward and reverse primers should be used in place of 6-FAM labeled forward primer and biotin labeled reverse primer, respectively.
2. Follow **Agarose Gels** directions above to ensure that aptamers are present without contamination (confirmed by a single, sharp band at the expected length for each sample).
3. Once sample purity has been confirmed by gel electrophoresis, mix PCR product with TOPO TA cloning kit reagents and incubate according to TOPO TA cloning kit directions<sup>7</sup> to generate plasmids.
4. Combine plasmids with One Shot chemically competent *E. coli* cells and heat shock according to directions included to transform cells.

- Transfer cells to selection plates, incubate, and inoculate colonies according to directions included to isolate individual colonies.
- Isolate plasmid DNA using PureLink plasmid miniprep kit according to PureLink plasmid miniprep kit directions.<sup>9</sup>
- Sequence isolated plasmid DNA by standard protocol, or send DNA for commercial sequencing. Ensure that final sample has been prepared in accordance with all of the commercial service's rules and restrictions.

### Aptamer Characterization

- Programs such as ClustalW identify conserved motifs in a pool of sequences. When using programs that identify motifs, it is necessary to remove the primer regions because the conservation of these regions will dominate the analysis.
- Programs such as *m*-fold predict the secondary structure of ssDNA and RNA molecules.
- Dissociation constants for individual aptamers can be determined using ACE as described for the bulk dissociation constants.

### Results

The flexibility of the CE-SELEX protocol allows it to be adapted to a wide range of target molecules. Coated capillaries can be used to eliminate adsorption of particularly difficult protein targets. Selection conditions and stringency are easily adjusted through modification of incubation and separation buffers. CE-SELEX has already been used for large proteins like Immunoglobulin E (IgE)<sup>1</sup> and HIV reverse transcriptase (HIVrt)<sup>5</sup>; small to medium proteins, including Human Epididymis protein 4 (HE4)<sup>9</sup> and Ricin<sup>2</sup>; the peptide Neuropeptide Y<sup>11</sup>; and even a small molecule, N-methyl mesoporphyrin IX (NMM)<sup>4</sup>. Below are example selections and dissociation constant measurements for IgE and NMM to illustrate the system's flexibility.

#### Selection

CE-SELEX is able to isolate aptamers for targets that induce minimal shifts in aptamer mobility. The best example of a "difficult" target due to its small size is NMM. At 580 Da, NMM only changes the aptamer mass by 2.3% upon binding, giving rise to a miniscule change in mobility. Because of this tiny change,  $t_{\text{det}}$  was chosen right at the unbound DNA peak's leading

edge (rather than 10-20 s prior), as illustrated in Figure 5. This adjustment of  $t_{\text{det}}$  illustrates one of the greatest strengths of CE-SELEX: flexibility. The stringency of partitioning can easily be adjusted from target to target and from round to round by changing  $t_{\text{det}}$  to collect more or less DNA. The adjustments were performed very accurately by monitoring the electropherogram and calculating  $t_{\text{cut}}$  in real time to account for any variability in migration.

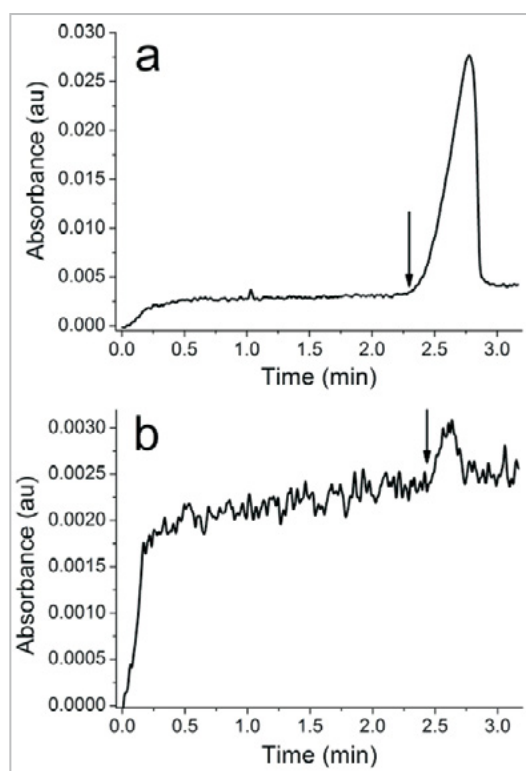


Figure 5: Selection electropherograms for (a) the first round of selection and (b) the second round of selection of NMM aptamers. Electropherograms for rounds 3-6 were similar to (b). Arrows indicate  $t_{\text{det}}$ , the end of the collection window. Detection: UV Absorbance @ 254 nm (Yang, J. Bowser, M. T., *Anal. Chem.*, 2013, 85 (3), 1525-1530.)



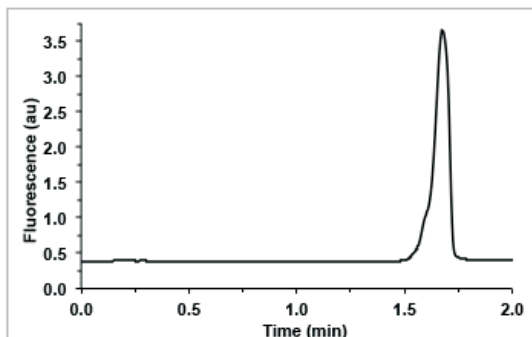


Figure 6: Selection electropherogram for the second round selection using LIF detection (Ex. = 488 nm; Em. = 520 nm). Electropherograms for rounds 3-5 were similar.

As described earlier, PCR yield limits the sample DNA concentration, and consequently the intensity of the DNA peak, in later selection rounds. For example, the DNA peak in Figure 5b is barely observable above baseline using UV absorbance detection. A slightly lower DNA concentration or noisier baseline might make the peak disappear altogether. Even if the peak is visible, the noise features can make it much more difficult to reliably determine where the DNA peak begins or ends. For these reasons, LIF detection is a preferable alternative to UV absorbance when possible. LIF elicits a much greater signal from fluorescently-labeled DNA than does absorbance, and far less noise is produced since the labeled DNA is typically the only fluorescent species in the sample. These factors together can produce a S/N  $10^2$ – $10^3$  times greater than absorbance detection. Figure 6 shows a representative second round collection electropherogram using LIF detection. A S/N of nearly 3,000 is achieved, compared to a S/N of ~5 using UV absorbance detection in Figure 5b. If a fluorescently-labeled aptamer is desired for later use (e.g. for imaging or binding characterization), isolation of aptamers from a labeled library is recommended. Addition of a fluorophore to an aptamer sequence post selection can potentially alter the aptamer's structure and decrease or eliminate its affinity for the target.

#### Dissociation Constant Measurements

An example of the dissociation constant measurement method described above is shown here for a particular IgE aptamer obtained by CE-SELEX.<sup>1,3</sup> Eight samples with IgE concentration from 0 nM to 150 nM, and with 25 nM of aptamer clone 4.6, were prepared and incubated at room temperature. Each

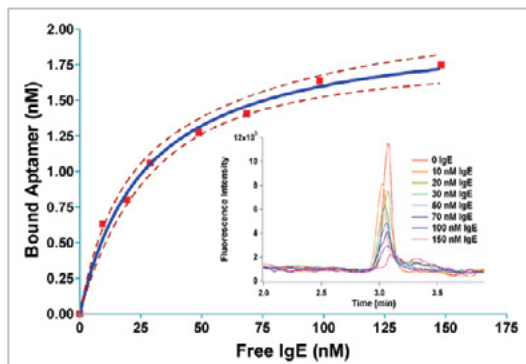


Figure 7: Binding curve for Clone 4.6 with results of rectangular hyperbola fit (nonlinear least squares). The dashed lines represent the 95% confidence band of the best fit curve. Inset: Electropherogram for each sample. Height of uncomplexed DNA peak decreases systematically from blank sample (no IgE) to 150 nM IgE sample. (Mendonça, S. D.; Bowser, M. T., *J. Am. Chem. Soc.* 2004, 126 (1), 20-21.)

sample was separated by CE (Fig. 7 inset) and the height of the unbound DNA peak was measured. The peak height decreased systematically with increasing IgE concentration as expected, down to about 10-15% of the blank height at 150 nM IgE. The bound fraction (left side of Eqn. 2) was then plotted against the sample IgE concentration, and the points were fitted using Eqn. 1 to estimate  $K_d$  (Fig. 7). The fit was performed on GraphPad Prism using least-squares regression analysis with a rectangular hyperbola. Clone 4.6 was estimated to have a  $K_d$  of  $27 \pm 8$  nM based on the fit. While an aptamer for IgE with slightly stronger affinity ( $K_d = 6$  nM) was obtained by conventional SELEX,<sup>10</sup> this aptamer required 15 selection rounds with negative selections, compared to four rounds without negative selection for Clone 4.6, demonstrating the efficiency of CE-SELEX.

#### Selectivity

In addition to high affinity, many aptamer sequences selected possess significant selectivity for the target of choice over other similar targets. In the case of NPY, aptamer sequences also had their affinity for human pancreatic peptide (hPP) quantified. hPP was chosen as a control because it has the same number of amino acids as NPY, shares 50% sequence homology, and acts upon some of the same receptors. After only four rounds, the six sequences tested at random had an average of 20-fold selectivity for NPY over hPP, with one sequence exhibiting 42-fold selectivity.<sup>11</sup>

p8

[Download this technical note](#)

Human IgE aptamers from 2-4 rounds of CE-SELEX were also incubated with human IgG and mouse IgE, both with significant sequence homology or sequence similarity to human IgE. None of the six sequences tested had any measurable affinity for human IgG (up to 20  $\mu$ M) or mouse IgE (up to 2  $\mu$ M). All sequences had better than 40 nM  $K_d$  for human IgE, indicating at least 500-fold and 50-fold selectivity over human IgG and mouse IgE, respectively.

### Conclusions

CE-SELEX has been used to successfully isolate aptamers for a variety of targets ranging from small molecules to large proteins and has made the selection process dramatically faster. CE-SELEX provides a number of benefits over traditional SELEX methods. The high resolving power and efficiency of CE generates high rates of enrichment allowing aptamers to be isolated after 2–4 rounds of CE-SELEX selection. Conventional selections typically require 8–15 selection cycles, a process that can take weeks to months to complete, and often generate weaker binding aptamers, as well as a large number of sequences that don't exhibit any affinity for the target. CE-SELEX performs selections in free solution, dramatically reducing the potential for nonspecific interactions with stationary surfaces and generating DNA pools in which nearly 100% of the sequences display target affinity. Additionally, free solution separation eliminates the difficulty of immobilizing the target molecule and doesn't require any knowledge of the target structure. CE-SELEX generates a large number of independent aptamer sequences, producing many lead compounds for therapeutic or diagnostic applications, and the incubation is highly customizable to generate aptamers that work well under any conditions. Finally, CE-SELEX is much more compatible with libraries that incorporate non-natural nucleic acids to improve aptamer stability. These non-natural nucleic acids often contain hydrophobic functional groups, which promote nonspecific interactions with hydrophobic filter materials, such as nitrocellulose, commonly used in conventional selections.

### Acknowledgment

Financial support for this work was provided by the National Institutes of Health (GM 063533) and the University of Minnesota.

### References

1. Mendonsa, S. D.; Bowser, M. T., In vitro selection of high-affinity DNA ligands for human IgE using capillary electrophoresis. *Analytical Chemistry* 2004, 76 (18), 5387-5392.
2. Tang, J. J.; Xie, J. W.; Shao, N. S.; Yan, Y., The DNA aptamers that specifically recognize ricin toxin are selected by two in vitro selection methods. *Electrophoresis* 2006, 27 (7), 1303-1311.
3. Mendonsa, S. D.; Bowser, M. T., In vitro evolution of functional DNA using capillary electrophoresis. *Journal of the American Chemical Society* 2004, 126 (1), 20-21.
4. Yang, J. B., M. T., Capillary Electrophoresis-SELEX Selection of Catalytic DNA Aptamers for a Small-Molecule Porphyrin Target. *Analytical Chemistry* 2013, 85 (3), 1525-1530.
5. Mosing, R. K.; Mendonsa, S. D.; Bowser, M. T., Capillary electrophoresis-SELEX selection of aptamers with affinity for HIV-1 reverse transcriptase. *Analytical Chemistry* 2005, 77 (19), 6107-6112.
6. Jing, M.; Bowser, M. T., Methods for measuring aptamer-protein equilibria: A review. *Analytica Chimica Acta* 2011, 686 (1-2), 9-18.
7. TOPO TA Cloning Kit for Sequencing, Invitrogen. [https://tools.thermofisher.com/content/sfs/manuals/topotaseq\\_man.pdf](https://tools.thermofisher.com/content/sfs/manuals/topotaseq_man.pdf) (accessed 24 May, 2016).
8. PureLink Quick Plasmid Miniprep Kits, Invitrogen. [https://tools.thermofisher.com/content/sfs/manuals/purelink\\_quick\\_plasmid\\_qrc.pdf](https://tools.thermofisher.com/content/sfs/manuals/purelink_quick_plasmid_qrc.pdf) (accessed 24 May, 2016).
9. Eaton, R. M.; Shallcross, J. A.; Mael, L. E.; Mears, K. S.; Minkoff, L.; Scoville, D. J.; Whelan, R. J. Selection of DNA aptamers for ovarian cancer biomarker HE4 using CE-SELEX and high-throughput sequencing. *Analytical and Bioanalytical Chemistry* 2015, 407 (23), 6965-6973.
10. Wiegand, T. W.; Williams, P. G.; Dreskin, S. C.; Jouvin M. H.; Kinet, J. P.; Tassed, D. Highaffinity oligonucleotide ligands to human IgE binding to Fc epsilon receptor I. *Journal of Immunology* 1996, 157, 221-230.
11. Mendonsa, S. D.; Bowser, M. T., In vitro selection of aptamers with affinity for neuropeptide Y using capillary electrophoresis. *Journal of the American Chemical Society* 2005, 127 (26), 9382-9383.

For Research Use Only. Not for Use in Diagnostic Procedures.

## Drug Discovery and Development



AB Sciex is doing business as SCIEX.

© 2016 AB Sciex. For research use only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of the AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

RUO-MKT-02-4213-A 05/2016



#### Headquarters

500 Old Connecticut Path, Framingham, MA 01701, USA  
Phone 508-383-7800  
[sciex.com](http://sciex.com)

#### International Sales

For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



# Nucleic acid sequencing

## Bacterial Artificial Chromosome (BAC) End Sequencing Analysis

Zhiming Jiang  
*Beckman Coulter, Inc.*

### Introduction

One of the goals of genetic analysis is to increase the understanding of gene function through the determination of DNA sequences that make up the genome of different species. In search of the best route to these ends, researchers have generated several different types of useful chromosomal maps. Eventually, the genome will be represented by DNA chromosome sequences with various levels of annotation. Bacterial artificial chromosomes (BAC), which typically contain 100 to 200 kb inserts of DNA, were designed as larger, more stable recombinant DNA clones that would represent the genome more uniformly than plasmid systems. Therefore, BACs are important tools for positional cloning, gene analysis, and physical mapping. During studies with BAC clones, it is often necessary to organize them into contigs. To finalize, join, and extend the contigs, both cloning and sequencing of the end of the inserts are required. However, BAC end sequencing has proven to be difficult due to the large molecular weight and secondary structure of the BACs. Here we present a new way of performing BAC end sequencing on the CEQ/GeXP Genetic Analysis System from Beckman Coulter with high degrees of accuracy, automation, and throughput. The same system could be applied to perform end sequencing of other large DNA constructs such as PAC, YAC, or cosmid, etc. This new method allows scientists to rapidly and efficiently sequence BAC DNA on a large scale, producing sequence data with relatively high signal levels, good quality values, and long read lengths.

### Methods

#### *Growth of BAC Clone and DNA Isolation*

A total of six human BAC clones (catalog number 96012) from Invitrogen were used for experiments. All clones were supplied as glycerol stocks and revived by transferring a small portion of the frozen sample onto an LB agar plate with Chloramphenicol (12.5 µg/mL) prior to incubation at 37° C for 12–16 hours. Single colonies were isolated from the freshly streaked plate and inoculated in a starter culture of 2-5 mL LB selective medium followed by an incubation of 12–16 hours at 37° C with vigorous shaking (~300 rpm). The starter culture was diluted into 500 mL selective LB medium (1/500 to 1/1000 dilution) and grew for 12–16 hours at 37° C with vigorous shaking. BAC DNA was then isolated using the QIAGEN Large-Construct Kit (catalog number 12462) with an expected yield of 20 to 50 µg BAC DNA from each 500 mL culture. To maintain fresh stocks of BAC clones, the BAC cultures were streaked onto a selective LB plate every other week and single colonies were isolated into selective LB medium. Glycerol stocks were then made by adding sterile glycerol into the BAC LB culture to achieve a final glycerol concentration of 20%. The glycerol stocks were then mixed thoroughly and frozen at -70° C.



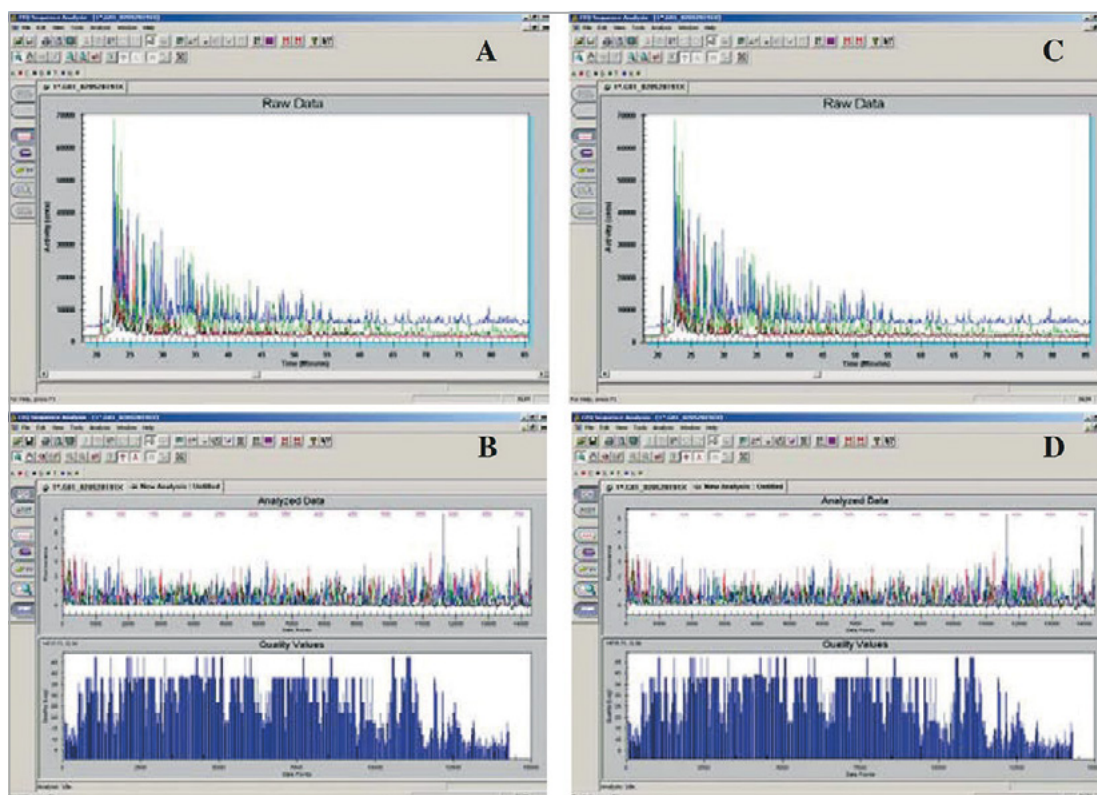
### End Sequencing Reaction Using DTCS Kit (P/N 608000) or DTCS

Quick-Start Kit (P/N 608120) BAC DNA was quantitated by UV spectroscopy at 260 nm. To confirm the size of each BAC clone, 1 µg of BAC DNA was digested by HindIII at 37° C for one hour and then separated on 1% agarose gel. All clones were shown to have a size of over 100 kb. BAC sequencing reactions were performed using DTCS kit or DTCS Quick-Start Kit (Tables 1 and 2). The cycling conditions are listed in Table 3. Sequencing reactions were purified by ethanol precipitation (DTCS Kit insert). Samples were dried in a speed vacuum, resuspended in 40 µL SLS, and separated on the GeXP Genetic Analysis System using the run method "LFR-a" (Table 4).

### Results

#### BAC Sequencing Analysis on the GeXP Genetic Analysis System

Raw data were collected and analyzed by the sequencing module of the GeXP software. All six clones were sequenced multiple times using the protocol described above. Figure 1 shows the typical sequencing results from BAC clones #1 and #6. Relatively high signal levels were achieved with high-quality value scores and long read lengths. Control sequences were obtained from GenBank and utilized as alignment templates. The 98% cutoff base number, a measurement of data accuracy, was calculated using the batch alignment function of the GeXP software. As shown in Table 5, using the recommended dye-terminator sequencing protocol and GeXP system, we were able to sequence all six BAC clones with high degrees of accuracy.



**Figure 1.** Typical BAC sequencing results with relatively high signal level, good quality values, and long read length. A: Raw data—clone 1. B: Analyzed data—clone 1. C: Raw data—clone 6. D: Analyzed data—clone 6.

p2

[Download this technical note](#)

Reagents	Amount
BAC DNA	2–3 µg
End Primer	60 pmoles
DTCS Premix	12.0 µL
Water	q.s.
Total	20.0 µL

**Table 1.** BAC Sequencing Protocol Using GeXP DTCS Kit.

Reagents	Amount
BAC DNA	2–3 µg
End Primer	60 pmoles
DTCS Premix	8.0 µL
Water	q.s.
Total	20.0 µL

**Table 2.** BAC Sequencing Protocol Using GeXP DTCS Quick-Start Kit.

Temperature	Time @ Temperature	Cycles
95° C	5 min	1
95° C	30 sec	50
50-55° C	10 sec	
60° C	4 min	
Hold at 4° C	Forever	1

**Table 3.** BAC Sequencing Cycling Conditions.

The overall average 98% cutoff base number from all six clones was 668 bp, whereas the overall average read length was 796 bp. Our end-sequencing success rate is approximately 95%.

## Discussion

To achieve the best BAC sequencing results, several issues need to be taken into consideration:

### 1) Fresh Stocks of BAC Clones

Most of the BAC clones contain 100 to 200 kb inserts of DNA. It is easy for the bacterial cells to lose large vector constructs when stored in selective liquid culture or on agar plates at low temperature. The percentage of BAC-carrying cells decreases dramatically after refrigeration for extended periods of time, with eventual loss of all BAC-containing clones. To avoid the loss of BAC, we recommend streaking the BAC culture onto a selective

<b>Capillary</b>	Temperature	50 °C
	Wait for Temperature	Yes
<b>Denature</b>	Temperature	90 °C
	Duration	120 seconds
<b>Pause</b>	Duration	0 seconds
<b>Injection</b>	Voltage	2.0 kv
	Duration	15 seconds
<b>Separate</b>	Stage 1	
	Primary Voltage	4.0 kv
	Ramp Duration	5 minute
	Stage 2	
	Separation Voltage	4.0 kv
	Start Time	5 minute
	Ramp Duration	0 minute
	Total Separation Duration	110 minutes

**Table 4.** BAC Sequencing Run Method "LFR-a".

LB plate every other week and inoculating single colonies into 2-5 mL selective LB medium prior to starting a large selective culture (500 mL). The solubility of chloramphenicol in water is much lower (2.5 mg/mL) than its solubility in ethanol (50 mg/mL). The recommended concentration for chloramphenicol in LB medium or agar is 12.5 µg/mL. However, the higher the concentration (up to 50 µg/mL) used, the more stringent growth conditions appeared to help BAC stability.

### 2) QIAGEN Large-Construct Kit

High-purity DNA is essential for achieving high-quality sequencing results. We suggested using the QIAGEN large-construct kit for BAC DNA isolation. This kit was designed for isolation of up to 50 µg ultrapure genomic DNA-free BAC, PAC, P1, or cosmid DNA. The recommended culture volume for BAC is 500 mL. Since large DNA constructs are generally present in low or very low copy numbers in cells, the actual yield varies

depending on the construct insert, the vector, and the host strain. The instructions in the kit should be followed closely.

In addition, the following suggestions are also helpful for obtaining pure BAC DNA with a higher yield:

- Reduce the culture volume to avoid overloading columns.
- Reduce DNA shearing by minimizing the vortexing or pipetting up and down during lysis.
- If experiencing difficulty in dissolving DNA, check whether the DNA was overdried or contaminated by isopropanol or salt.
- If the BAC DNA pellet is invisible, rinse the bottom of the centrifuge tube with 0.5 mL of water, transfer the solution to a 1.5 mL tube, and then dry the DNA using a speed vacuum. Avoid over-drying the pellet. Resuspend the BAC DNA in 10–20  $\mu$ L of TE buffer (pH 8.0).
- Try to dissolve BAC DNA in as low a volume of buffer as possible. The concentration of BAC DNA should be kept at 0.3–1.0  $\mu$ g/ $\mu$ L.

### 3) Sequencing Primer

Good primer design is critical for achieving good sequencing results. We recommend using a commercial software package to design sequencing primers. In this study, we used the Oligo 6 software package from Molecular Biology Insights, Inc. Dye-terminator reactions for plasmid vectors usually require 3.2 pmol of primer per reaction. For an equivalent amount of BAC DNA, there are 50–100 times fewer priming sites available. In addition, the BAC target contains increased numbers of imperfect binding

sites which may titrate away some primer molecules. Both longer primers and higher primer concentrations seem to be helpful for sequencing quality improvement. Out of the series of primer concentrations we tested, the best working concentration for BAC sequencing is 60 pmol primer/20  $\mu$ L sequencing reaction (Tables 1 and 2).

### 4) BAC Template

High-molecular-weight DNA may be hard to dissolve at high concentration. Over-drying of BAC DNA pellets may make this problem worse. Warming the solution slightly, and allowing more time for the pellet to dissolve can help overcome this problem. We performed sequencing reactions using different amounts of template DNA per reaction. The higher amount of BAC DNA in the reaction, the better the sequencing results. However, using too much BAC DNA may result in a sticky cluster in the sequencing reaction, incomplete denaturing of the template, or insufficient primer annealing, which therefore leads to lower signal levels. We recommend using 2.5–3.0  $\mu$ g BAC DNA in a 20  $\mu$ L sequencing reaction.

### 5) Cycling Conditions

Different combinations of denaturing, annealing, and extension temperature and time were tested. The optimal conditions are presented in Table 3. The initial 5-minute incubation step at 95° C is included to insure complete denaturation of the BAC templates prior to primer annealing. Preheating has proven to improve sequencing quality as well.

Clone Number	ResGen Clone ID	Human Clone Number	Number of Reactions	Average Read Length (bp)	Average 98% Cutoff Base Number
1	2015B8	Chr 1 clone RP11-122M14	11	804	689
2	2326G7	Chr 7 clone CTA-281B9	14	821	674
3	2581C3	Chr 15 clone CTD-2033D15	8	695	521
4	2049P16	Clone hRPK.60_A_1	9	780	732
5	2337M5	Chr 17 clone CTD1-2246P4	10	856	640
6	2253H2	Clone hRPK.60_A_1	11	817	756
				Overall Average 796	Overall Average 668

**Table 5.** Sequencing Data Summary on All Six BAC Clones.

### 6) DNA Precipitation

Following the dye terminator sequencing reactions, excess free dyes were removed by ethanol precipitation. To achieve good DNA precipitation and avoid losing pellet, we recommend:

- Add stop solution and glycogen as a master mix prior to adding cold ethanol.
- Mix thoroughly before adding the cold 95% ethanol.
- Add cold 95% ethanol to the tube, mix thoroughly, then centrifuge immediately at 14,000 rpm, 4° C for 20 minutes.
- Rinse the DNA pellet twice with 500 µL (instead of 200 µL) cold 70% ethanol followed by immediate centrifugation.
- Avoid over-drying the DNA pellet.

### 7) Separation on the GeXP

Good resuspension of the final pellet in SLS is also critical for good sequencing results from BAC DNA. Add 40 µL of SLS to each DNA pellet and incubate at room temperature for 5 minutes. Vortex for 10 seconds and then pipette up and down multiple times to insure complete resuspension in SLS. Reactions were loaded to the GeXP and separated by "LFR-a" method with a separation voltage of 4.0 kV for 110 minutes.

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. A-1936-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701, USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





# Improved Sequencing of Plasmids on the SCIEX Genetic Analysis System

## *By a Simple Template Preheating Procedure*

Mark Dobbs, Ph.D.  
*Beckman Coulter, Inc.*

Recent advances in automated sequencing technology have centered around capillary-based systems. Such systems are highly automated and simple to run. One potential drawback, however, concerns the ability of such systems to sequence large or otherwise difficult plasmid DNA samples. Here we report a simple, cost-effective and convenient solution to overcome problems associated with sequencing plasmid DNA on the SCIEX CEQ/GeXP Genetic Analysis System (GeXP).

### Methods

Two different plasmid DNA templates were used in this study: a GC-rich cDNA cloned into pUC18 (3.5 Kb), and large DNA plasmid (>12 Kb).

Plasmid DNA templates were prepared using commercially available kits from either QIAGEN (QIAprep<sup>®</sup>) or Promega (Wizard<sup>®</sup>). Plasmids were prepared according to manufacturers' instructions.

Templates were quantitated by running restriction digested aliquots on agarose gels and comparing ethidium bromide staining intensities to known standards.

Samples were diluted in deionized H<sub>2</sub>O, the tubes were closed and then heated for 1 minute at 96° C in an MJ Research PTC. After 1 minute at 96° C, the temperature was ramped down immediately to 25° C, and the samples were used in sequencing reactions.

Sequencing reactions were set up as described in the kit insert 608118AE DTCS sequencing protocol. Cycling conditions were as described: 96° C for 20 seconds, 50° C for 20 seconds, 60° C for 4 minutes.

Post-reaction clean up was by ethanol precipitation as described in the DTCS sequencing protocol 608118AE.

The sequencing reaction products were run on the GeXP. The capillary temperature was set to 40° C and the separation voltage to 8.2 kV; otherwise all parameters were as for the default DTCS - 2 method. Data analysis was carried out using SCIEX genetic analysis sequencing software.

### Results

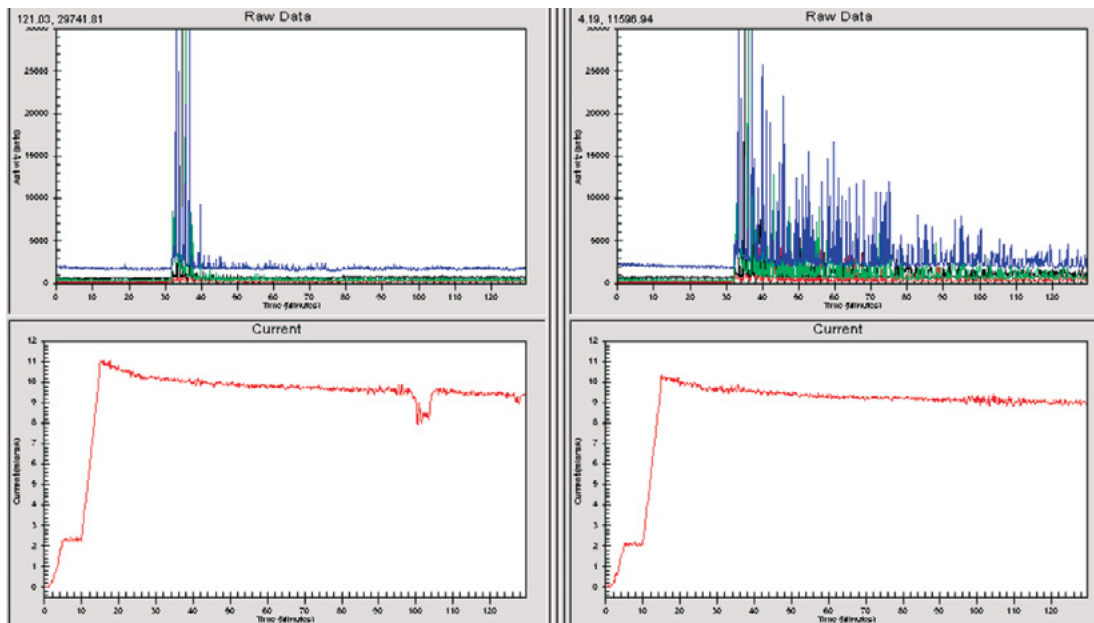
The GeXP allows the user to monitor the raw data and the current in each individual capillary. These two features allow the user to quickly diagnose problems associated with sequencing templates on the GeXP system.

Low signals can lead to inaccuracies in base calling as well as failure to automatically call bases by the GeXP. In the example shown in Figure 1a, low signal strength led to no sequence being called. A simple incubation of the plasmid sample resuspended in water just before adding the cycle sequencing reaction mix (described in the Methods section) results in a substantial increase in the raw data signal. Figure 1b shows the signal improvement when the same plasmid as used in the experiment from Figure 1a is treated at 96° C for 1 minute prior to sequencing. This treatment improved the sequence call from no bases to 704 bases with the first 637 bases being called at 98% accuracy.

The data shown in Figure 2 compares the same GC-rich plasmid prepared by a different method (Promega Wizard). In this case, the preheating treatment improves the signal strength and the current stability. Unstable currents can lead to anomalous peak spacing and hence errors in the sequence call. In the example shown here, the pretreated sample called 536 bases and the sequence was 99.3% accurate. The untreated sample was only 65% accurate at 550 bases.

p1

[Download this technical note](#)



**Figure 1a (left).** Raw data and current profiles for GC-rich DNA prepared using QIAGEN QIAprep column. 100 fmol (230 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 30,000 counts and the current profile is scaled to 12  $\mu$ A.

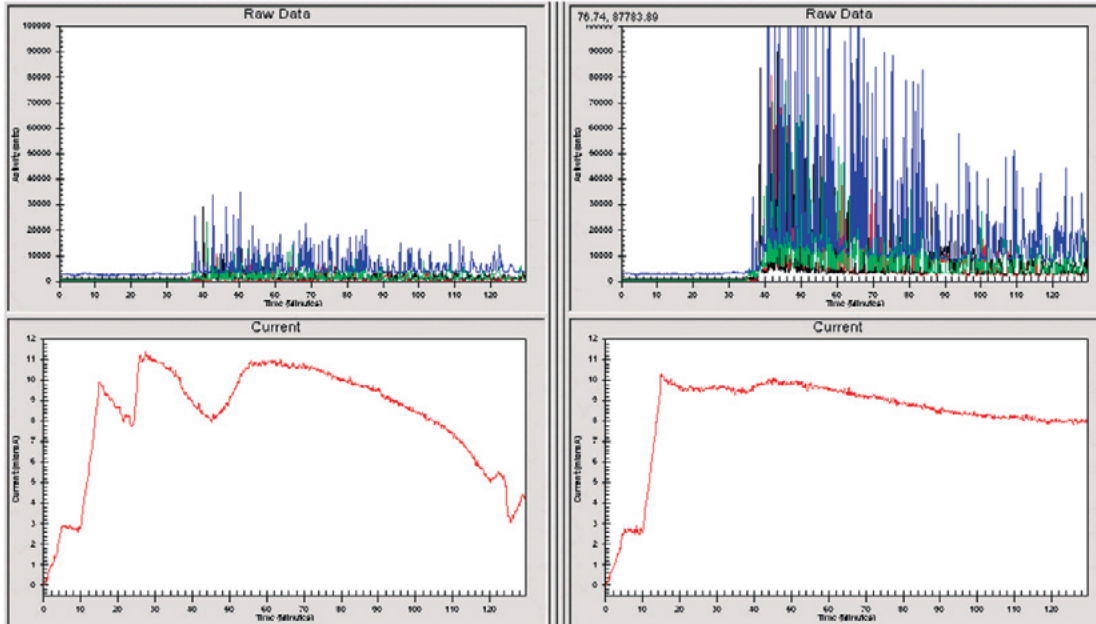
**Figure 1b (right).** Raw data and current profiles for GC-rich DNA prepared using QIAGEN QIAprep column and pre-treated for 1 minute at 96° C. 100 fmol (230 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 30,000 counts and the current profile is scaled to 12  $\mu$ A.

Analysis of the data from a large plasmid template (shown in Figure 3) showed that the treated sample called a total of 579 bases, whereas before treatment only 461 bases were called. Although both sequences were accurate to at least 98% over the length called, the more stable current led to a longer sequence being called.

## Conclusions

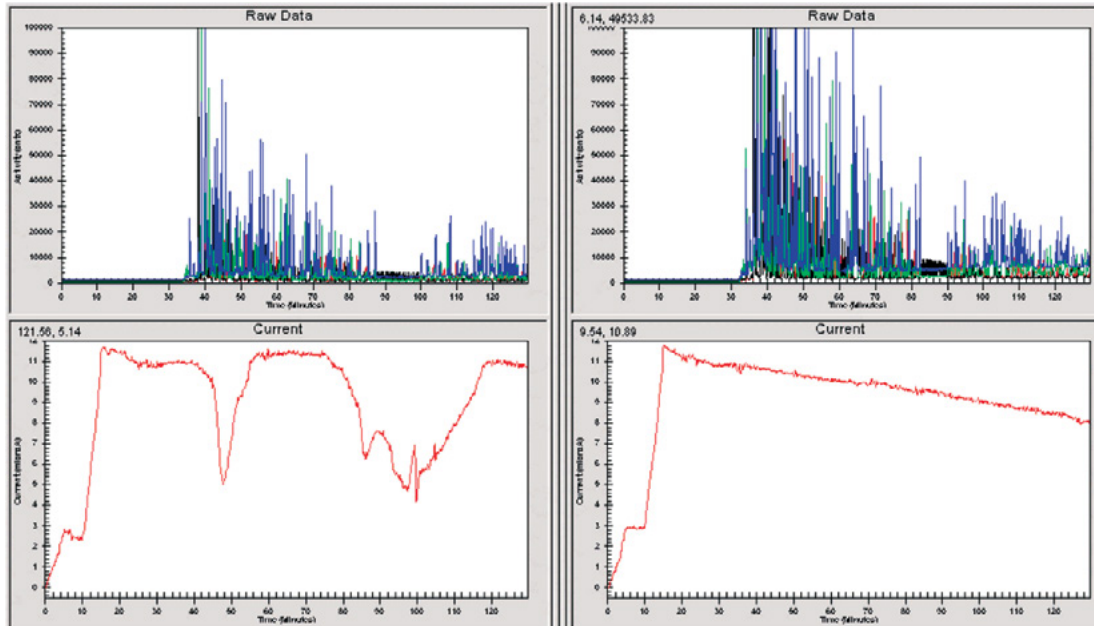
We have demonstrated a simple, cost-effective, and convenient method for improving automated sequencing of plasmid DNA templates on a capillary-based system. The preheating procedure

is carried out with the samples diluted in deionized H<sub>2</sub>O prior to adding the remaining sequencing reaction components. In a few cases, a 1-minute treatment at 96° C can lead to decreased signal strength for longer sequencing fragments. In these cases, preheating the template at 86° C for 5 minutes produced results similar to the ones shown here (data not shown). Still other templates required preheating at 96° C for 3 minutes to stabilize current profiles. It appears that 96° C for 1 minute is a good compromise treatment. In certain cases, the treatment may have to be altered as described above in order to maximize the observed signal and current benefit. Future work will investigate the mechanism of action in this simple procedure.



**Figure 2a (left).** Raw data and current profiles for GC-rich DNA prepared using Promega Wizard column. 100 fmol (230 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 30,000 counts and the current profile is scaled to 12  $\mu$ A.

**Figure 2b (right).** Raw data and current profiles for GC-rich DNA prepared using Promega Wizard column and pre-treated for 1 minute at 96° C. 100 fmol (230 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 30,000 counts and the current profile is scaled to 12  $\mu$ A.



**Figure 3a (left).** Raw data and current profiles for a large (12 Kb) plasmid prepared using QIAGEN QIAprep column. 100 fmol (800 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 100,000 counts and the current profile is scaled to 12  $\mu$ A.

**Figure 3b (right).** Raw data and current profiles for a large (12 Kb) plasmid prepared using QIAGEN QIAprep column and pretreated for 1 minute at 96°C. 100 fmol (800 ng) of DNA was used in the sequencing reaction. Raw data is scaled to 100,000 counts and the current profile is scaled to 12  $\mu$ A.





## Method for Long Sequencing Read Lengths on the SCIEX Genetic Analysis System

Doni Clark, Jim Thorn, and Keith Roby  
*Beckman Coulter, Inc.*

### Introduction

Automated DNA sequence analyzers have increased the throughput and ease of use of an application that has become a core technique for many areas of genetic analysis. For large de novo sequencing projects, automated DNA sequence analyzers have obviously shown their advantage and strength. While most DNA sequencing is confirmatory in nature, with required read lengths of 700 bases or less, some researchers need to maximize base read lengths to economize on laboratory resources.

Although the CEQ/GeXP Genetic Analysis System produces sequence read lengths greater than 700 bases using the standard sequencing methods provided with the instrument, longer sequencing read lengths are possible by modifying the separation parameters. This bulletin presents a strategy for increasing base read lengths by 20% or more by modifying the separation voltage, duration, capillary temperature, and injection time parameters as compared to the existing CEQ LFR1 sequencing separation method. Applying this new separation method can achieve average 98% accuracy base read length cutoffs greater than 900 bases with some individual samples obtaining 98% accuracy cutoffs of 1000+ bases.

### Method

#### DNA Purification

Overnight cultures of DH5 cells containing pUC18 as a control plasmid and pUC19 vector containing a 1.2 kB Glucouronidase gene insert (pGus) were grown overnight in rich bacterial media with 100 µg/mL ampicillin to an optical density of 4-5. The cultures were transferred to deep square-well plates at 1 mL per well. Plasmid DNA preparation was performed on a Biomek® FX Laboratory Automation Workstation from Beckman Coulter using both the Promega Wizard® SV 96 Plasmid DNA Purification

System and the Qiagen QIAprep® 96 Turbo Miniprep plasmid DNA purification kit. The Biomek FX program for the Promega Wizard has been described previously in Beckman Coulter Application Bulletin A-1907A.<sup>1</sup> This Biomek FX program was modified for use with the Qiagen QIAprep DNA purification kit.

#### DNA Sequence Reaction

Since many DNA prep technologies yield a large quantity of supercoiled plasmid which can affect the performance of sequence fragment separations, appropriate preheat treatment of plasmid DNA for any particular cell type and purification method combination should be determined by performing a temperature-versus-time matrix.<sup>2</sup> In addition, for obtaining longer read length separations, the preheat treatment should be as gentle as possible in order to relax the supercoiled plasmids while not nicking them so much that producing longer fragments during the cycle sequencing reactions becomes problematic. Both of the DH5-grown plasmids in this study were preheat treatment optimized using a matrix of three temperatures (65° C, 76° C, and 86° C) and four time lengths (1 to 4 minutes). The final preheat treatment condition for both of these purification methods that maintained appropriate signal strength for the longer sequencing fragments was 3 minutes at 86° C.

Sequencing reactions were performed with both the pGus plasmid and the pUC18 control plasmid using the 608120 DTCS Quick Start Kit. All reactions used 70 fmol of plasmid DNA template which was thermal cycled as suggested.<sup>2</sup> While this quantity of DNA template worked well in our studies, many researchers may need to adjust this template quantity for their specific needs based on template type and size. Another means of increasing the overall signal and hence the amount of longer fragment signal is to increase the number of cycles in the thermal cycling parameters. This can be combined with resuspension of the final sequencing products in less Sample

p1

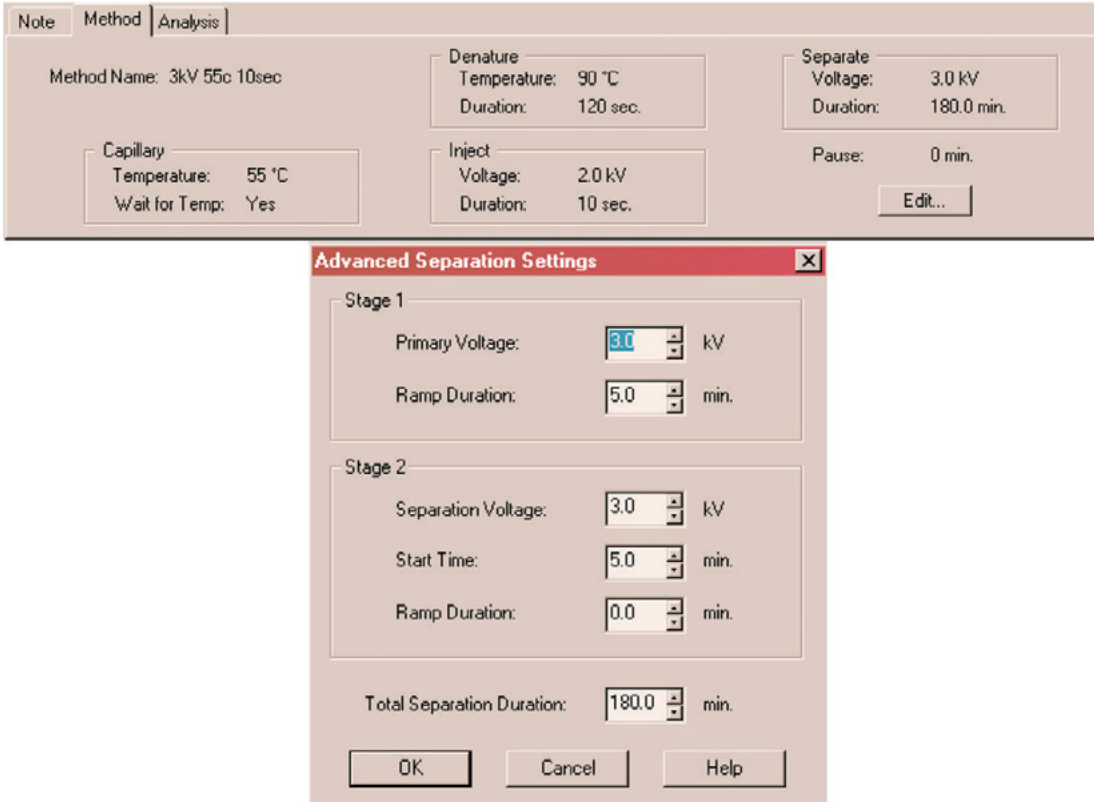
[Download this technical note](#)

Loading Solution (30  $\mu$ L instead of 40  $\mu$ L). It should be noted that incrementally increasing signal strength by increasing the effective concentration of the sequence fragments in the sample causes more sample to be injected onto the capillary which negatively affects the resolution of the longer fragments. Long fragment signal strength and resolution are both critical factors for achieving long base read lengths; therefore, a balance must be achieved that provides adequate signal strength for detection but not so much as to negatively impact resolution. Purification of the sequencing reaction products was performed by the ethanol precipitation technique.<sup>3</sup> After drying the plates, the sequencing products were resuspended in 40  $\mu$ L Sample Loading Solution, overlaid with mineral oil, and loaded onto the CEQ/GeXP Genetic Analysis System.

## Results and Discussion

To obtain longer sequencing read lengths, the separation method parameters were modified to reduce the voltage and increase the data collection duration over a range of capillary temperatures. After performing a series of voltage versus temperature experiments, it was found that the following separation method yielded the most consistent results (Figure 1).

Using the new separation method parameters, a comparison was performed against the default LFR-1 separation method supplied with the GeXP. For the pUC18 DNA prepared with the Promega kit, 24 individual samples were run. All other sample sets for both the control pUC18 and pGUS DNAs represent a total of 48 individual samples per DNA purification chemistry and separation method. Comparison of peak resolution is performed in Figure 2. Visually, peak resolution is similar between the



The image shows two screenshots from a software interface. The top screenshot is the 'Method' tab of the 'Sample Setup' module. It displays the following parameters:

- Method Name: 3kV 55c 10sec
- Capillary Temperature: 55 °C
- Wait for Temp: Yes
- Denature Temperature: 90 °C
- Denature Duration: 120 sec.
- Inject Voltage: 2.0 kV
- Inject Duration: 10 sec.
- Separate Voltage: 3.0 kV
- Separate Duration: 180.0 min.
- Pause: 0 min.
- An 'Edit...' button is located at the bottom right.

The bottom screenshot is the 'Advanced Separation Settings' dialog box, which is open over the 'Separate' section of the method tab. It contains the following settings:

- Stage 1:**
  - Primary Voltage: 3.0 kV
  - Ramp Duration: 5.0 min.
- Stage 2:**
  - Separation Voltage: 3.0 kV
  - Start Time: 5.0 min.
  - Ramp Duration: 0.0 min.
- Total Separation Duration:** 180.0 min.
- Buttons: OK, Cancel, Help

Figure 1. Sample Setup module Method tab.

p2

[Download this technical note](#)

standard LFR1 separation method and the new longer read method until approximately base 400. Beyond that point, the new longer read separation method clearly demonstrates improved peak resolution.

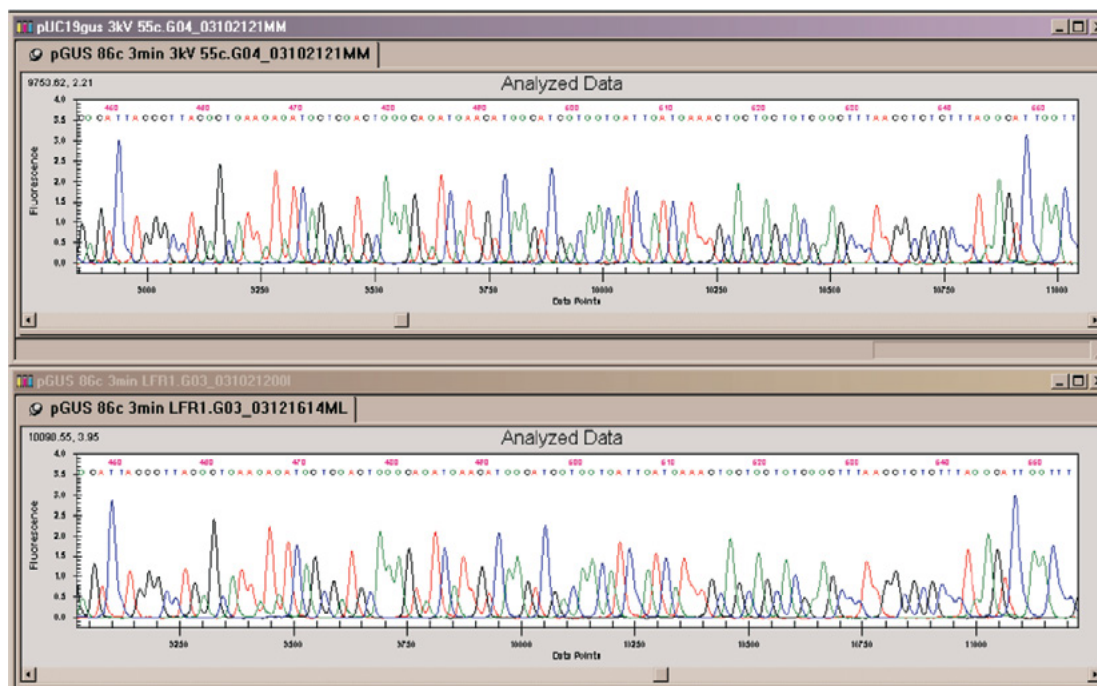
*Note: Capillary lifetime studies were not tested with the new, longer-read separation method.*

Table 1A shows representative base read length 98% accuracy cutoffs for each plasmid, separation method, and purification chemistry combination. The sample names are coded as such: QpUC18 and QpGUS represent the template DNAs purified by the Qiagen chemistry, while pUC18 and pGUS from Table 1B are the DNA templates purified by the Promega chemistry. 86c 3min indicates the preheat treatment used for all samples, and LFR1 or 3kV 55c indicate, respectively, the standard GeXP sequencing method or an abbreviation of the new optimized separation method.

The 98% accuracy cutoff results were averaged and compared for each set of templates (Table 2). In each combination of DNA

purification chemistry and plasmid template, the new separation method demonstrated an increase of greater than 20% for the base read length 98% accuracy cutoff values.

Use of this separation method increases the overall base sequence yield from a single reaction by approximately 20% and thereby maximizes the amount of sequence information obtained from a single reaction. In some instances where DNA templates may not yield sufficient fluorescent signal, it may be helpful to increase either the number of cycles during the sequencing reaction thermal cycling, or to increase the injection time duration in the separation method. Be aware, though that adding too much DNA template to the sequencing reaction to boost signal strength by making more product or increasing the injection time too much will greatly impact the resolution of the longer sequencing fragments and should be avoided. This is an important point to consider since, in order to achieve the longest possible sequencing read lengths, it is absolutely critical to maintain resolution quality.



**Figure 2.** Comparison of separation methods for pGUS demonstrating improvement in resolution of longer read separation method at bases 510 to 600.

[Download this technical note](#)

Qiagen QIAprep 96 Turbo Miniprep Results Name	Total Read Length	98.00% Cutoff Base #	Qiagen QIAprep 96 Turbo Miniprep Results Name	Total Read Length	98.00% Cutoff Base #
QpUC18 86c 3min 3kV 55c.A07_03110103KY	1042	972	QpGUS 86c 3min 3kV 55c.A09_03110608V9	1037	906
QpUC18 86c 3min 3kV 55c.B07_03110103KY	1083	996	QpGUS 86c 3min 3kV 55c.B09_03110608V9	905	905
QpUC18 86c 3min 3kV 55c.C07_03110103KY	1055	974	QpGUS 86c 3min 3kV 55c.C09_03110608V9	900	900
QpUC18 86c 3min 3kV 55c.D07_03110103KY	1061	1005	QpGUS 86c 3min 3kV 55c.D09_03110608V9	1063	961
QpUC18 86c 3min 3kV 55c.E07_03110103KY	1271	1032	QpGUS 86c 3min 3kV 55c.E09_03110608V9	1012	911
QpUC18 86c 3min 3kV 55c.F07_03110103KY	1050	1020	QpGUS 86c 3min 3kV 55c.F09_03110608V9	1033	906
QpUC18 86c 3min 3kV 55c.G07_03110103KY	1115	1026	QpGUS 86c 3min 3kV 55c.G09_03110608V9	969	950
QpUC18 86c 3min 3kV 55c.H07_03110103KY	997	936	QpGUS 86c 3min 3kV 55c.H09_03110608V9	1022	912
QpUC18 86c 3min L FR1.A10_03111209JW	788	788	QpGUS 86c 3min LFR1.A12_03110720XN	770	768
QpUC18 86c 3min LFR1.B10_03111209JY	803	803	QpGUS 86c 3min LFR1.B12_03110720XN	790	761
QpUC18 86c 3min LFR1.C10_03111209K0	798	796	QpGUS 86c 3min LFR1.C12_03110720XN	788	767
QpUC18 86c 3min LFR1.D10_03111209K1	807	807	QpGUS 86c 3min LFR1.D12_03110720XN	778	778
QpUC18 86c 3min LFR1. E10_03111209K3	803	796	QpGUS 86c 3min LFR1. E12_03110720XN	744	744
QpUC18 86c 3min LFR1.F10_03111209K5	804	801	QpGUS 86c 3min LFR1.F12_03110720XN	791	766
QpUC18 86c 3min LFR1.G10_03111209K7	806	795	QpGUS 86c 3min LFR1.G12_03110720XN	798	761
QpUC18 86c 3min LFR1.H10_03111209K9	810	810	QpGUS 86c 3min LFR1.H12_03110720XN	711	711

**Table 1a.** Qiagen QIAprep Base Read Length 98% Accuracy Cutoff.

Qiagen QIAprep 96 Turbo Miniprep Results Name	Total Read Length	98.00% Cutoff Base #	Qiagen QIAprep 96 Turbo Miniprep Results Name	Total Read Length	98.00% Cutoff Base #
pUC18 86c 3min 3kV 55c.H05_03103001P3	1059	1025	pUC19gus 3kV 55c.A04_03102121MM	1169	959
pUC18 86c 3min 3kV 55c.B04_03102922FY	1075	961	pUC19gus 3kV 55c.B04_03102121MM	1029	943
pUC18 86c 3min 3kV 55c.C04_03102922FY	1025	1004	pUC19gus 3kV 55c.C04_03102121MM	1060	977
pUC18 86c 3min 3kV 55c.D04_03102922FY	1029	1023	pUC19gus 3kV 55c.D04_03102121MM	1100	985
pUC18 86c 3min 3kV 55c.E04_03102922FY	1054	996	pUC19gus 3kV 55c.E04_03102121MM	1015	971
pUC18 86c 3min 3kV 55c.F04_03102922FY	1049	1022	pUC19gus 3kV 55c.F04_03102121MM	1077	981
pUC18 86c 3min 3kV 55c.G04_03102922FY	1124	982	pUC19gus 3kV 55c.G04_03102121MM	1104	1017
pUC18 86c 3min 3kV 55c.H04_03102922FY	1108	1030	pUC19gus 3kV 55c.H04_03102121MM	1071	1044
pUC18 86c 3min LFR1.A03_03102920TH	742	730	Gus 8643 40 µL LFR1.A05_031018054A	769	769
pUC18 86c 3min LFR1.B03_03102920TH	823	769	Gus 8643 40 µL LFR1.B05_031018054A	833	812
pUC18 86c 3min LFR1.C03_03102920TH	830	801	Gus 8643 40 µL LFR1.C05_031018054A	811	811
pUC18 86c 3min LFR1.D03_03102920TH	826	800	Gus 8643 40 µL LFR1.D05_031018054A	847	837
pUC18 86c 3min LFR1.E03_03102920TH	802	789	Gus 8643 40 µL LFR1.E05_031018054A	833	809
pUC18 86c 3min LFR1.F03_03102920TH	818	801	Gus 8643 40 µL LFR1.F05_031018054A	821	821
pUC18 86c 3min LFR1.G03_03102920TH	807	786	Gus 8643 40 µL LFR1.G05_031018054A	799	789
pUC18 86c 3min LFR1.H03_03102920TH	803	783	Gus 8643 40 µL LFR1.H05_031018054A	796	796

**Table 1b.** Promega SV96 Base Read Length 98% Accuracy Cutoff.



### References

1. Quick Start Guide to Purifying Plasmids on the Biomek® FX Using Beckman Coulter Wizard SV 96 Reagents. *Beckman Coulter Application Information Bulletin Quick Start Guide A-1907-QS* (2001).
2. DTCS kit insert 608118 AG
3. Roby, K., Gull, H. A Rapid and Efficient Method for the Post-Reaction Clean Up of Labeled Dye Terminator Sequencing Products. *Beckman Coulter Application Information Bulletin A-1903A* (2001).

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. T-1975-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701, USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



# SNP analysis & genotyping

## SNPs. Mutations and DNA Sequence Variation Analysis Using the GeXP SNPStart Kit

Nitin Udar, Jana Mariana, Margaret Porter and Doni Clark  
*Beckman Coulter, Inc.*

### Introduction

SNP analysis is routinely used for linkage, association studies, pharmacogenetics, forensics, etc. Similarly, mutation screening requires detection of single base changes, as well as analysis of insertions, deletions and translocations among other molecular events. In this study we have used the single-base extension (SBE) technology to extend an unlabeled primer flanking a target site with a fluorescent labeled terminator. The product is separated and detected on a capillary electrophoresis platform – GeXP Genetic Analysis System. All the reagents necessary for the SBE reaction including the fluorescent dyes, enzyme and buffer, are bundled in a single ready-to-run SNPStart master mix (Figure 1). The master mix is optimized to generate accurate as well as well-balanced allele signals in a SNP reaction. We demonstrate multiple levels of plexing capability using SNP panels from one to 10-plex. In addition, multiplex Cystic Fibrosis (CF) mutation panels were also tested using patient DNA samples. Signal balance between two heterozygous alleles is important for accurate detection. We measured this signal balance for each of the two alleles of the six possible heterozygote combinations and found an average signal ratio of 1:1.6. The allele calling accuracy for a 10-plex reaction is >99%. Comparisons were done using different sources of DNA, including DNA extracted from cell lines and prepared from whole genome amplification. These were used as templates for PCR and subsequently used for primer extension reactions. Our results demonstrate an efficient process of producing more than 960 SNP/mutation screens per plate. DNA extracted on the Vidiera™ NsP nucleic acid sample preparation platform (based upon the Biomek® NX platform) was tested for the SNP application. This testing was carried out with the intent of increasing process performance and reducing possible sampling errors and hands on time per sample. In this report we demonstrate the capabilities and ease of use of the SNPStart Primer Extension kit.

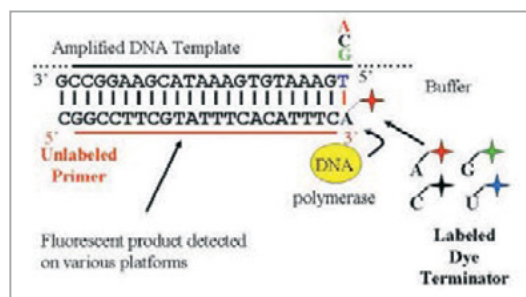


Figure 1. Single-Base Extension Reaction.

### Methods and Materials

#### PCR Primer Design

Primers for PCR were designed using the Primer 3 software developed by the Whitehead Institute for Biomedical Research. PCR primers were designed with melting temperature between 60°C - 80°C with optimum between 70°C - 75°C. The length of the primer was designed between 23-32 nt with an optimum of 27 nt. Avoid designing PCR primers in repeat regions.

The desired length of the PCR products was between 150-1000 bps. The lower limit is important when columns are used for PCR purification. Most commonly used columns have a lower limit of approximately 100 bp. PCR products larger than 1000 bps may need an additional denaturation step (94° C for 1 min) immediately prior to primer extension cycling. Smaller PCR products are desired because they will have a less likelihood of nonspecific binding of other products. For a multiplex PCR situation, design the product lengths so that they do not overlap by less than 30bp. This enables visualization of each individual band if run on a regular agarose gel to verify successful PCR before further processing the sample. In these experiments we used the following PCR cycling conditions:

94° C/2 min followed by 35 cycles of 94° C/30 sec, 60° C for 1 min and 72° C for 1 min (the time will need to be increased further for longer PCR products) and then held at 4° C. The PCR products have to be further processed to remove all the excess primer and dNTP before proceeding with the SNP reaction. This can be achieved by either treating the PCR product with EXO-SAP (GE HealthCare, US) or column purification. Follow manufacturer's recommendation for this step. PCR products were electrophoresed either on an agarose gel or on an Agilent Labchip (Agilent Technologies, US) to verify specificity and make sure there were no artifact bands.

Allele balance is partially dependent on the amount of template used in the PCR reaction. If the starting DNA concentration is too little it will lead to a successful PCR reaction but the allele balance could be unacceptable. To avoid this problem of allele bias introduced during PCR, add sufficient template DNA (in the range of 100-200 ng).

## Interrogation Primers

Interrogation primers were designed using the same considerations as mentioned above (for PCR primer design). Interrogation primers can sometimes be difficult to design using the ideal parameters because the locus is fixed. There are only two choices in designing the primers – either on the forward strand or the reverse strand. Vary the number of bases in the primer to increase the T<sub>m</sub>. Hairpin and other secondary structures are difficult to avoid in the interrogation primer. Select the one that has the least secondary structure. It is sometimes easier to design primer on both strands of the SNP and selecting the one that works the best. In a multiplex panel, a poly A tail should be added to separate the different SNPs. The poly A tail is added to the 3' end of the primer. In a panel with more than 7 primers, the interrogation primers should be spaced out by 6 nts. For panels 7-plex and less, space out the primer by at least 8 nt. Interrogation primer can be designed on either strand of the target sequence. Even if primers are spaced out at fixed intervals expect a few variations in the mobility.

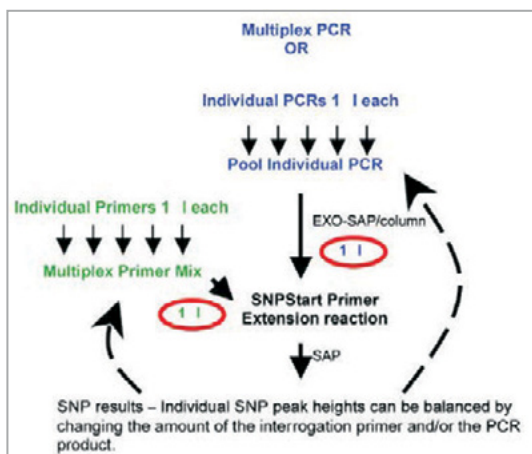
Primer	Size	Base Change	Difference in Size	Sequence
SNP1-U43	25	G/T	N.A.	AAAAGTGAAAACCTAAGGGAACTGC
SNP2-U36	31	A/T	6	AAAAAACTGCGGGATTCTCGTTGGCAATGAA
SNP3-U13	37	C/G	6	AAAAAAAAAAAAAGTGTGGATCTTGTGAATAAATGAG
SNP4-U19	43	C/A	6	AAAAAAAAAAAAAAAAAAAAACCCATGACTTCTGCCCCACAC-CCTG
SNP5-U30	49	C/T	6	AAAAAAAAAAAAAAAAAAAAAAAAAATCTCAAAGGAAA-CACTTTAGCCCA
SNP6-U30	55	G/A	6	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGAGATGCAACATCAGGTGATAAG
SNP7-U4	61	C/A	6	AAATTTT-GAACTCTCCTAACAGAGCTGA
SNP8-U5	67	G/A	6	AAGACAGCCTCGAGACTCAGGTGA
SNP9-U34	72	C/G	5	AA AAAAATCTAATGGCATGTGATACTCATCC
SNP10-U37	77	C/T	5	AA AAAAAAAACCTCTGTGGACCACATCTGAGTCCTG

**Table 1.** Details of interrogation primers used in SNP panels.

[Download this technical note](#)

Primer	Size	Base Change	Sequence
A455E	23	g>t	ctgctccagtgatccagcaacc
G1349D	35	g>a	AAAAAAAAAAAAAAtggggctgtgtcctaagccatg
N1303K	47	g>c	AAAAAAAAAAAAAAAAAAtgatcactcactgttcataggatccaa
R1162X	53	g>a	AAAAAAAAAAAAAAAAAAAAAAAAAAtgtggcatgtcaatgaacttaagactc
2789+5G	59	c>t	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAacacaatctacacaataggacatggaata
W1282X	64	g>a	AAAgtctgggattcaataacttgcaacagtg
R117H	73	g>a	AActtctatgaccocggataa-caaggaggaac
F508	37	c>t	AAAAAAAAtgccttgcaccattaaagaaaatcat
3849+10kb	31	c>a	gagcttccatctgttcagtattaaatgg

**Table 2.** Details of interrogation primers used in Cystic Fibrosis panels.



**Figure 2.** Multiplex Setup.

Generally you would require a few optimization steps to balance the signal intensity in a multiplex SNP reaction. The protocol to follow is as follows: Carry out the multiplex or single plex PCR as per your protocol. Separate it on a agarose gel to make sure there are no artifact bands (overload the product a little bit to be sure you can visualize the artifacts if any). Make sure all excess primer is removed from this reaction before proceeding with the SNP primer extension reaction. A clean and robust PCR product is the single most important factor in determining the success of the SNP reaction. Resuspend interrogation primers to a final

concentration of 100 μM. Mix 1 μl of each of the individual primers in a tube and use 1 μl of this mixture for the SNP reaction (Figure 2). Follow the same rule of thumb for the PCR products if doing a poolplex (where each of the individual PCR products after ExoSAP/Column purification is mixed in a 1:1 proportion – take 1 μl of this mixture for the SNP primer extension reaction).

### SNP Reaction

Interrogation primers used in the SNP panel are as mentioned in Table 1 and for Cystic Fibrosis Panel in Table 2. Template concentration in the various SNP reactions are shown in Table 3. If the PCR reaction required DMSO or other denaturing agents in the reaction – purify the PCR reaction using column purification to remove the denaturing agent. If using such PCR products for the SNP reaction add Betaine in the range of 0.5 - 1.5 M to obtain robust primer extension products. The protocol as described in Table 4 was followed for the SNP and Cystic Fibrosis panels described in this report. It is important to note that the master mix has glycerol as one of its components. Therefore pipette the master mix with care. Protect the master mix as well as the reaction tubes from exposure to light.

After mixing all components of the SNP reaction the tubes were given a short spin to consolidate the mixture to the bottom of the tube. This was followed by thermal cycling the tubes at 94° C for 20 seconds followed by 45° C for 10 seconds. The single-plex reactions were cycled 25 times and multiplex reactions 36 times followed by holding at 4° C.



SNP reaction	Individual PCR-Template concentration in the mixture	Individual Primer concentration in the mixture	Dilutions	Loading volume
10-plex	0.3 - 3.5 ngs	0.4 - 2.0 $\mu$ M	1:3	1.0 $\mu$ l
2-plex	1.4 ngs	0.1 - 1.0 $\mu$ M	1:6	1.0 $\mu$ l
Single-plex	7.0 -40.0 ngs	0.1 - 1.0 $\mu$ M	1:10 - 1:20	1.0 $\mu$ l

**Table 3.** Concentration of template and primer used in the SNP reaction.

Component	Volume
SNPStart Master Mix	4.0 $\mu$ L
DNase RNase free PCR grade wate	x.x $\mu$ L
Interrogation Primer (0.1 - 1.0 $\mu$ M)	x.x $\mu$ L
PCR product or template (1-100 fmoles)	x.x $\mu$ L
Total Volume	10.0 $\mu$ L

**Table 4.**

After thermal cycling, the excess dye terminators in the SNP reaction were cleaned up using 0.25 U of Shrimp Alkaline Phosphatase (SAP), 1  $\mu$ l of SAP buffer and 1.75  $\mu$ l of DNase RNase free water. Load 1  $\mu$ l of the treated SNP reaction into 39.5  $\mu$ l of SLS and 0.5  $\mu$ l of Size Standard 80 (Table 4). Overlay the samples with mineral oil. After loading the sample plate and buffer plate on to the GenomeLab/GeXP Genetic Analysis System, select SNP-1 as the run method in the sample plate layout. Before starting the run make sure the capillaries are preheated to 50° C immediately prior to starting the run. This ensures proper separation for the first row of samples. Preheating is not required for subsequent rows within the same run. Observe the raw data for the first run by opening the results in the sequence analysis software. The sample peak heights should ideally be 50%-150% the peak height of size standard 13. Make appropriate dilutions of the cleaned SNP product in SLS if signal is too high. In such situations the dilution step can be avoided by decreasing the number of cycles for the primer extension reaction. Individual SNP peaks should be evenly balanced by increasing or decreasing their volume in the primer mix. In cases with a mixture of homozygous and heterozygous peaks the heterozygous peak heights should be adjusted to approximately half the height of the homozygous peaks. If necessary use a higher (100  $\mu$ M) or lower (1  $\mu$ M) starting primer concentration for peaks that need to be adjusted. Ideally the

peak heights should be between 8,000-80,000 rfu. Peaks with saturated signal will lead to errors in data analysis including missing peaks.

### Data Analysis

For Software version 8.0 and 9.0 users, install the SNP Software Update before data analysis. The software needs to be installed only once. To set up analysis parameters for the SNPStart data, use the fragment analysis module and follow the steps listed below.

1. Create a new study by selecting the sample raw data that needs to be analyzed, then select "Next".
2. Select "Default SNP Analysis Parameters" from the drop-down menu. The following parameters should be modified to correctly analyze the SNPStart data.
3. Click "Edit", then from the General Tab, change the Slope Threshold to 25 and Relative Peak Height Threshold to 15. These values can be adjusted, depending on the raw data to be analyzed.
4. Select the "Advanced" Tab and then choose the "SNP ver. 2" Dye Mobility Calibration.
5. Click "Save As", and then name the new Analysis method with a unique name, e.g. "SNPStart- Analysis Parameters." This method is now set up to analyze the data generated using the GenomeLab™ SNPStart Primer Extension Kit. This method can be further modified for adjusting the threshold or if SNP locus tags are added to the study. SNP locus tags can be created by clicking on the "SNP Locus Tag" tab in the analysis parameters section. Select "New locus" and fill in the relevant information. The text entered for the "Locus tag" will appear in the locus tag list. The text entered for the "Locus name" will appear in the analyzed results above the appropriate allele peaks. The "Apparent Fragment Size" should be determined by taking the average peak of previously analyzed results. Remove alleles that are not present in the SNP in "Allele ID's for" section. For example in an A/T SNP remove the C and G allele by deleting them. Save these settings with a different name.

Once created, the new analysis parameters will be available for subsequent projects, provided that they are in the “default” project within the same database. The parameters can also be selected for automatic analysis following each run. For frequently used parameter sets, copy and paste the new analysis parameters into the User Template Database in Data Manager to automatically populate all newly created databases with the analysis parameters.

To confirm validity of analyzed results: a) Compare the raw data with the analyzed results. b) Verify that the number of peaks (not cross-talk) is the same in raw and analyzed data. This does not take into account possible noise peaks which typically should be small. c) During a single separation, the ratio of the peak heights in the raw data within a SNP product should remain the same in the analyzed data. If there are errors in analysis due to dye matrix estimation, reanalyze the data using system dye spectra.

### Genome Amplification

The amount of DNA used for testing can become a constraint for SNP applications where a large number of different SNPs are tested or in situations where very little DNA is available. Whole genome amplification provides a method to increase the amount of starting material that can be used for downstream testing.

A total of 20 ng of DNA were amplified using GenomiPhi<sup>®</sup> DNA Amplification Kit (GE Healthcare) and used to amplify eight Cystic Fibrosis mutations. An aliquot of 0.5 µl of the amplified product was used for the subsequent locus specific PCR reaction in a total volume of 10 µl. After ExoSap treatment, a 0.5 µl aliquot of the PCR product was used as template in the final SNP reaction as described above. After SAP digestion, a 1.0 µl of the product along with 0.5 µl of SS-80 and 38.5 µl of SLS was aliquoted into a well. The products were separated on the GeXP system as described above. Ideally the peaks should be within 50%-150% the peak height of size standard 13.

### Automated Platform

Beckman Coulter Biomek Automation platform (based upon the Biomek NX platform) was used to purify genomic DNA from lymphoblastoid cell lines. This DNA was used as a template for PCR and the products obtained were subsequently used as templates for the SNP reaction.

### Results and Discussion

We carried out single-plex and multiplex SNP reactions using a panel of 10 SNPs. PCR products were obtained for each individual SNP as separate reactions. Primers were designed so that none of the products exceeded 1000 bps in size. During design they were also masked for repeats using the human repeat database. PCR products were verified on a agarose gel or Agilent chip to make sure there were no artifact bands. If artifact bands were present the PCR conditions were altered to obtain a clean single band. A clean PCR product is one of the most important criteria that determines the outcome or results of the SNP reaction. The PCR products from individual reactions for a set of predetermined heterozygous individuals were mixed together in a 1:1 ratio. Similarly SNP interrogation primers (100 µM conc.) were also mixed in a 1:1 ratio. From each of these mixtures, a total of 1 µl each was used for the SNP reaction. The individual template concentrations in the 10-plex template mixture were in the range of 0.3 to 3.5 ng with an average of 1.2 ng. The individual primer concentrations in the 10-plex primer mixture were in the range of 0.4 to 2.0 µM with an average of 1.0 µM. As illustrated in Figure 3, balanced allele signal was obtained using 3 separate SNPs with a different heterozygote combination (C/A, G/T and A/T).

During data analysis, a dye spectra is generated based on the raw data. Sometimes there is not enough data in an individual sample run for the software to calculate the necessary dye spectra. In such situations, a system dye spectra can be used to analyze the results.

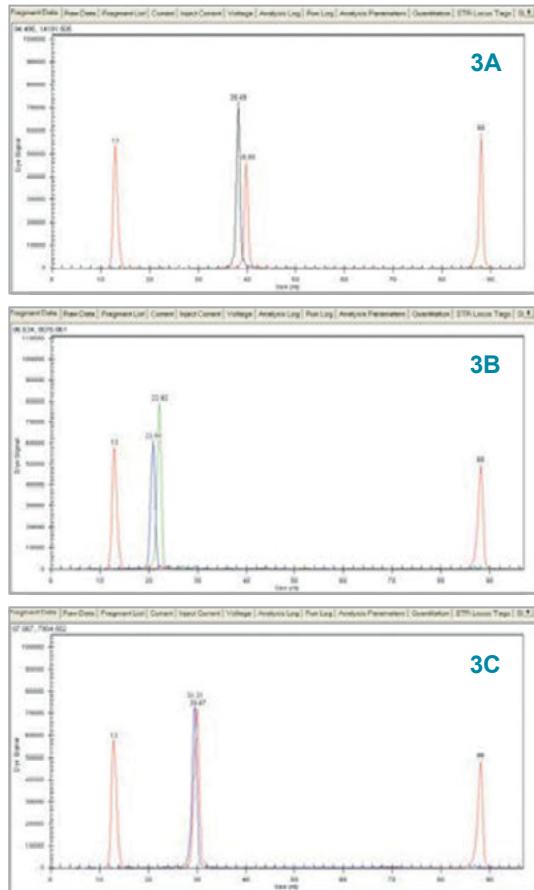
Figure 4 shows a 2-plex reaction that was analyzed using a system dye spectra. Slope threshold and relative peak height threshold can be changed based on the assay to obtain fewer identification of nonspecific peaks. Figures 4 and 5 illustrate that if the slope threshold is increased from 30 to 90 and the relative peak height threshold is increased from 20 to 50, the shoulder peaks in the second C/A SNP are not called anymore.

PCR products were obtained for 10 different SNP's from a set of known heterozygous DNA samples. These 10 individual products were mixed and tested using a panel of 10 SNP interrogation primers. Figure 6 illustrates the data obtained from the primer extension reaction using the SNPStart master mix.

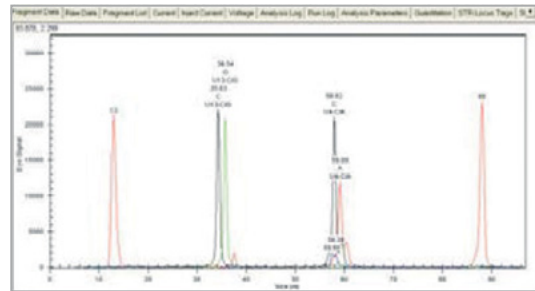
All ten loci are heterozygous and illustrate an extreme example but in fact the allele heights for each individual peak are well balanced. Using a full plate of data (96 X 10-plex), we determined the average ratio of the signal balance to be 1:1.6 (1:1 is a ideal theoretical balance). Locus sequence characteristics influence the balance for individual SNP's. Figure 6B shows the same results but locus tags were also used during analysis. Selecting only the possible genotypes and canceling the other bases is

important during locus tag set up. For example, when setting up a C/A locus tag, remove the G and T options in the possible genotype. This also acts as a screening tool during allele assignment. This helps to filter out nonspecific peaks of the other colors. Use locus tags only in samples that actually will have those peaks. Unused locus tags will create unnecessary entries in the fragment list.

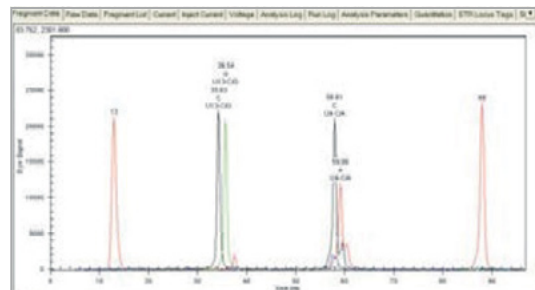
Similar multiplex panels were also tested to screen a panel of Cystic Fibrosis mutations. Two different panels were tested on different patient DNA samples. Figure 7 and Figure 8 shows the results obtained for these experiments. All four patients' DNA showed the correct genotype and mutations.



**Figure 3.** Single-plex SNP reactions - 3A) Heterozygous C/A alleles. 3B) Heterozygous G/T alleles-1C) Heterozygous A/T alleles.



**Figure 4.** Two-plex SNP reactions - both heterozygous C/G and C/A alleles. The two-plex SNP data is analyzed using A slope threshold of 30 and relative height threshold of 20. Small shoulder peaks are observed for the C/A SNP. These peaks are more commonly observed in PAGE purified primers as compared to HPLC purified interrogation primers. System Dye spectra was used to analyze the data as compared to Calculated dye spectra in order to get the correct SNP allele assignments.

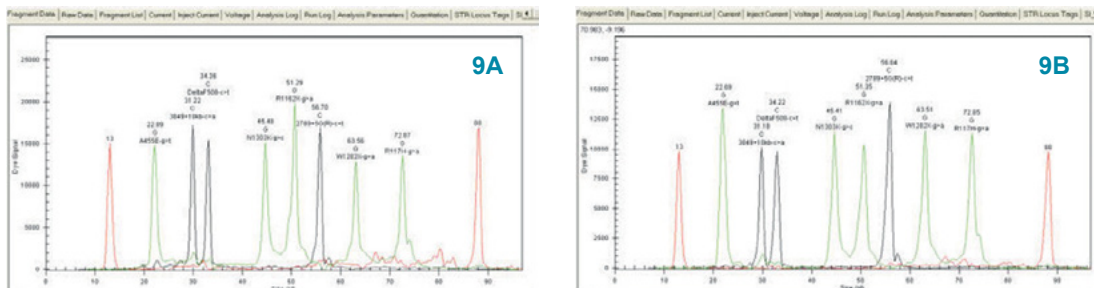


**Figure 5.** Two-plex SNP reactions - both heterozygous C/G and C/A alleles. The two-plex SNP data is analyzed using A slope threshold of 90 and relative height threshold of 50. Note the shoulder peaks on the C/A SNP are not called as in Figure 4.

p6

[Download this technical note](#)





**Figure 9.** SNPStart data for results obtained using 2 DNA samples purified on a Vidiera NsP platform (based upon the Biomek NX platform) tested for the Cystic Fibrosis Panel of mutations.

SNP genotyping has become a significantly powerful tool for research. SNP frequencies in the genome makes them ideal for testing for many different applications. However each individual SNP is only marginally informative (compared to microsatellite markers). To compensate for the informativeness a large number of SNPs are need to be screened. Although this may seem easy to do, the amount of starting DNA sample may make this a limiting endeavor. Genome amplification using phi polymerase offers a useful alternative. We tested genome-amplified DNA as starting material to find out if all the loci were equally amplified and if there was any difference in allele bias due the extra amplification step. Figure 8 shows that there were no significant differences in the results obtained using genome-amplified DNA for the Cystic Fibrosis locus as compared to genomic DNA isolated from lymphoblastoid cell lines.

For high throughput screening or to automate the screening process we also tested DNA samples purified from

lymphoblastoid cell line using the Vidiera™ NsP nucleic acid sample preparation platform (based upon the Biomek NX platform). DNA extracted on this platform gave robust SNP results illustrating that this platform can easily be combined to generate an automated workflow in a laboratory for accurate and efficient processing of samples (Figure 9).

## Summary

In summary, the GenomeLab™ SNPStart Kit allows users to obtain multiplex SNP data from templates generated from different sources of DNA. The color balance and allele signals are optimized in this kit. The kit comes as a ready-to-use single tube reagent that is based on a simple primer extension technology with no additional secondary steps involved. The kit has been demonstrated to run 10-plex SNP reactions. Therefore, a total of 960 genotyping results can be accomplished in a 96-well plate format.





## Multiplex SNP Analysis: Screening Factor V R506Q (Leiden) Mutations

Zhiming Jiang,<sup>1</sup> Milhan Telatar,<sup>2</sup> Wayne W. Grody<sup>2</sup>

<sup>1</sup> Beckman Coulter, Inc.

<sup>2</sup> University of California, Los Angeles

### Introduction

Single nucleotide polymorphisms (SNPs) are major contributors to genetic variation, making up approximately 80% of all known polymorphisms, and their density in the human genome is estimated to be, on average, one per 1000 base pairs. Although SNPs are mostly biallelic (less informative than short tandem repeats), they are more frequent and mutationally stable, making them suitable for association studies in which linkage disequilibrium (LD) between markers and an unknown variant is used to map disease-causing mutations. In addition, because SNPs have only two alleles, they can be genotyped by a simple yes or no assay rather than a length measurement, making them more amenable to automation.

### SNP Analysis

The SNP analysis involves three stages: SNP discovery, validation, and scoring. SNP discovery involves identifying potential SNPs by technologies such as sequencing, SSCP, or DHPLC. SNP validation confirms the real mutations and their allele frequencies in different populations. The SNP scoring phase includes genetic analysis such as association or linkage. The technological and economic goal is accurate, easy, cheap, and fast large-scale SNP genotyping.

### Primer Extension

The single-base primer extension is a method of sequencing the precise location of a SNP site. It utilizes the inherent accuracy of DNA polymerase to determine the presence or absence of the specific nucleotide at the SNP site. A specially synthesized DNA primer is used to anneal to the SNP site of interest. The primer

anneals one base short of the target SNP. DNA polymerase inserts the complementary dideoxy nucleotide terminator to the SNP site (Figure 1). This technology provides a simple multiplexable SNP genotyping solution with high accuracy and reproducibility, which is applicable to multiple platforms including capillary electrophoresis, mass spectroscopy, flow cytometry, HPLC, microarray, etc. The CEQ/GeXP Genetic Analysis System, CEQ DNA Size Standard 80, and SNP-Primer Extension Kit were newly developed to provide users an accurate, simple, cheap, and robust solution for SNP scoring and validation based on single-base primer extension technology.

### Factor V R506Q (Leiden) Mutation

Screening for genetic variants that predispose individuals or their offspring to disease may be performed at the general population level. Factor V R506Q (Leiden), causing activated protein C (APC) resistance, was discovered in 1994 and is the most common genetic risk factor for venous thrombosis. It is present in 5% of Caucasian Americans, 20% of idiopathic first venous thrombosis cases, and 60% of venous thrombosis cases in pregnant women. In this application note, we demonstrate that the single-base primer extension and GeXP system provided a novel way to screen this factor V mutation in the human population with a high degree of throughput, automation, and accuracy. The same method could be designed and applied to validate or score thousands of other mutations associated with human disease in research labs.

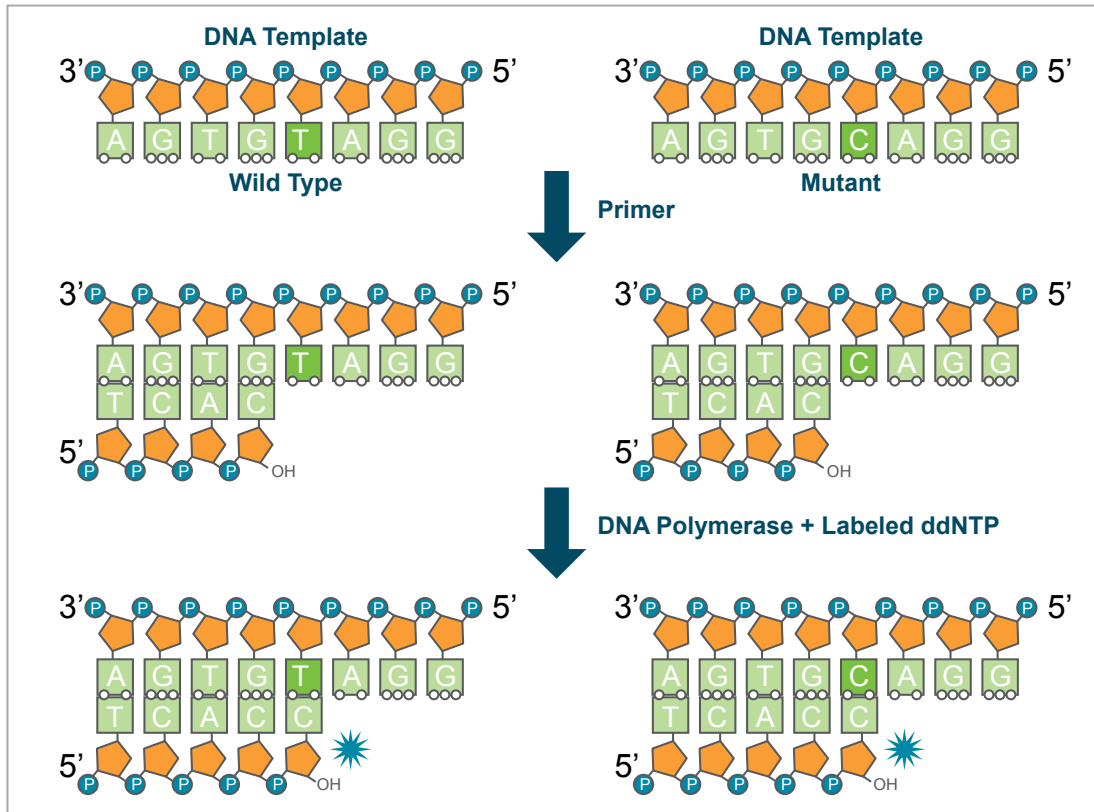


Figure 1. Primer extension theory.

## Methods and Results

### DNA Size Standard 80 (P/N 608397)

The CEQ Size Standard 80 was developed specifically for analysis of small fragments on fluorescent detection systems. The DNA Size Standard Kit-80 contains two reference fragments labeled with the D1 (red) dye for the use in sizing SNP fragments in the approximate size range 20 to 80 nts. It is designed to accommodate a wide range of sizes for up to twelve multiplex or poolplex SNP fragments, and, therefore, provides enhanced flexibility in choosing the length of SNP primers. We recommend the customer to vortex size standard 80 well to achieve equal injection of the two fragments before loading on the GeXP.

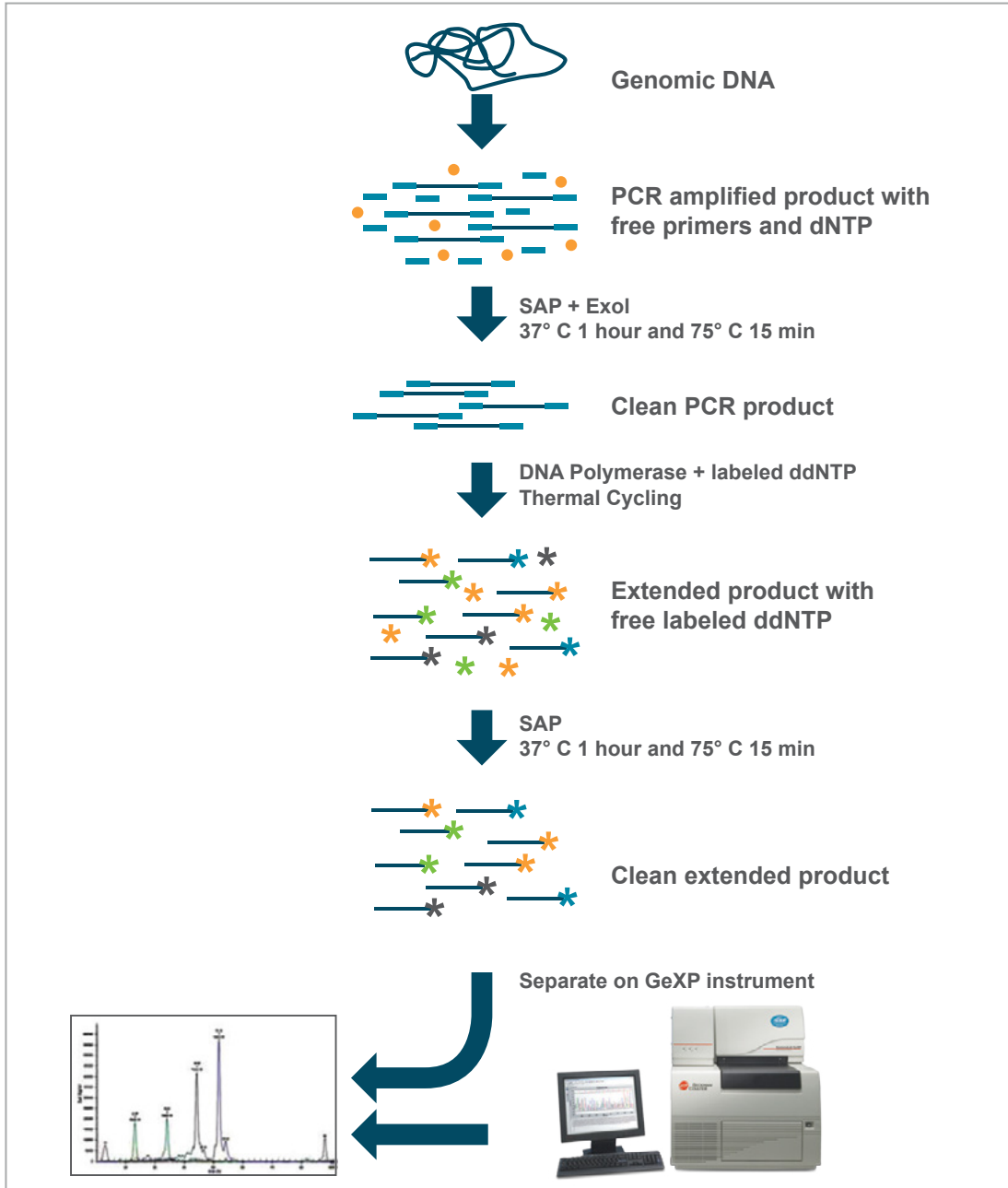
### SNP-Primer Extension Kit (P/N 390280)

As shown in Figure 2, the primer extension assay contains three major steps: template cleanup, singlebase primer extension, and extended product cleanup.

1. **Template cleanup:** Both plasmids and PCR\* products may serve as templates for primer extension. Because PCR products normally contain residual primers and free dNTP, which would lead to generation of non-specific products, it is highly recommended to clean up PCR templates prior to primer extension. The simplest method is to digest the single-strand primers and free dNTP by incubating with Shrimp Alkaline Phosphatase (SAP) and Exonuclease I (Exo I) at 37° C for 1 hour and then inactivate the enzymes at 75° C for 15 min. Both of the enzymes may be added directly into the PCR products without any dilution in a ratio of 6 µL PCR product: 2 U SAP: 1 U Exo I.

p2

Download this technical note



**Figure 2.** Overview of primer extension assay and data collection by GeXP Genetic Analysis System.

[Download this technical note](#)

2. **Primer extension reaction\***: The single-base primer extension utilizes the annealing of a regular synthesized DNA primer one base short of the SNP site to DNA template. In the presence of dye-labeled ddNTP, DNA polymerase will add the complementary dye terminator to the 3' end of DNA primer corresponding to SNP site. The primer annealing and extension reactions are normally thermocycled to achieve high signal levels from the extended products. The protocol for setting up primer extension reactions with the CEQ™ SNP-Primer Extension Kit is shown below in Table 2. For multiplex reactions, reduce water accordingly to accommodate all primers and templates. The thermal cycling conditions are shown in Table 3. The annealing temperature is adjustable to 5 degrees below the lowest primer  $T_m$ .

Component	Volume (μl)
10X Reaction Buffer	210
ddUTP Dye Terminator	210
ddGTP Dye Terminator	210
ddCTP Dye Terminator	210
ddATP Dye Terminator	210
Polymerase Enzyme	105
<b>Total</b>	<b>1155</b>

**Table 1.** Prepare SNP-Primer Extension Premix Using GeXP SNP-Primer Extension Kit.

Temperature	Volume (μl)
SNP-Primer Extension Premix	11.0
SNP primer(s) 1.0 μM	1.0
Template(s) 100 nM	1.0
Water	7.0
<b>Total</b>	<b>20.0</b>

**Table 2.** Primer Extension Protocol Using GeXP SNP-Primer Extension Kit.

Temperature	Time @ Temperature	Cycles
96° C	10 sec	25
50° C	5 sec	
72° C	30 sec	
Hold at 4° C	Forever	1

**Table 3.** Thermal Cycling Conditions.

3. **Cleanup extended products**: The free dye terminators frequently co-migrate with extended product during capillary electrophoresis, which leads to the generation of nonspecific signals. To avoid this interference, each extended product needs to be incubated with 1 unit of SAP at 37° C for 1 hour and then 75° C for 15 min. to remove the 5' phosphoryl groups of the free-labeled ddNTP prior to loading on the GeXP.

We recommend loading 0.5 μL of the reaction samples with SLS and size standard as a starting point (Table 4). If signal over-ranging is observed during analysis, it may be necessary to dilute the samples (1:40) with SLS prior to adding to the above mixture. Cover each sample with one drop of light mineral oil and run the plate on the GeXP using the SNP-1 separation method (Table 5).

Reagent	μL / Reaction
SLS*	39.0
Size Standard 80	0.5
Extension Reaction Samples	0.5

**Table 4.** Loading Mix.

		SNP-1
<b>Capillary</b>	Temperature	50 °C
	Wait for Temperature	Yes
<b>Denature</b>	Temperature	90 °C
	Duration	60 seconds
<b>Pause</b>	Duration	0 seconds
<b>Injection</b>	Voltage	2.0 kv
	Duration	15 seconds
<b>Separate</b>	Stage 1	
	Primary Voltage	6.0 kv
	Ramp Duration	1.0 minute
	Stage 2	
	Separation Voltage	6.0 kv
	Start Time	1.0 minute
	Ramp Duration	0.0 minute
	Total Separation Duration	16 minutes

**Table 5.** Run Method: SNP-1.



## GeXP Genetic Analysis System

The GeXP Genetic Analysis System enables the capillary electrophoresis system to perform automated sizing and allele calling from single or multiplexed SNP products. The new size standard 80 provides approximate sizing for SNP locus identification. A new fragment-sizing algorithm was developed to handle multiplexed SNP products. The GeXP software utilizes a mechanism similar to that for STR loci to analyze SNP. An overview of performing SNP analysis by GeXP software is shown in Figure 3. The raw data files are analyzed under editable analysis parameters developed specifically for SNP analysis, including SNP Locus Tags, SNP Dye Mobility Correction, Size Standard-80, etc. Background peaks may be excluded by modifying the slope threshold. The analyzed results display loci and alleles in the fragment list. No more than one fragment labeled with a given dye may belong to the same locus, while up to four peaks may be labeled as alleles of the same locus, but each must be labeled with a different dye. Dye mobility corrections are applied during data analysis. Multiplexing may be achieved by spacing between adjacent SNP products by 4 to 8 nts.

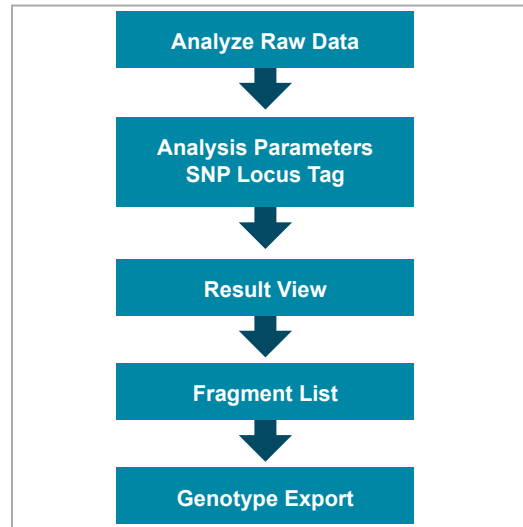


Figure 3. SNP analysis flow chart using GeXP software.

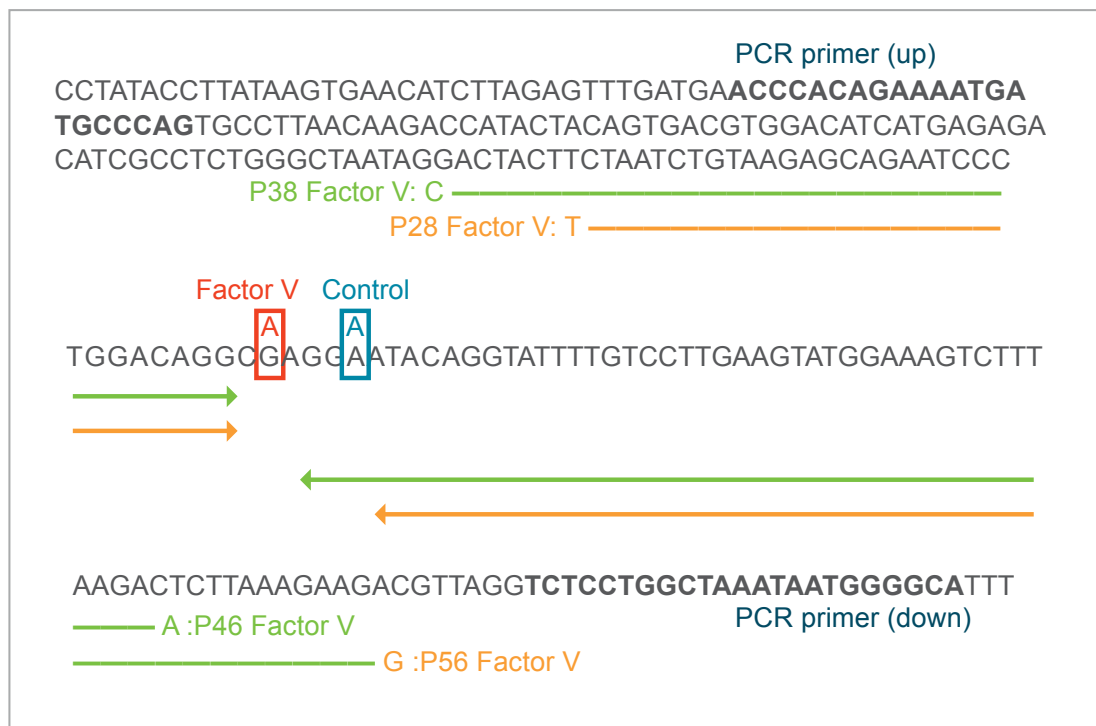


Figure 4. The factor V mutation locus.

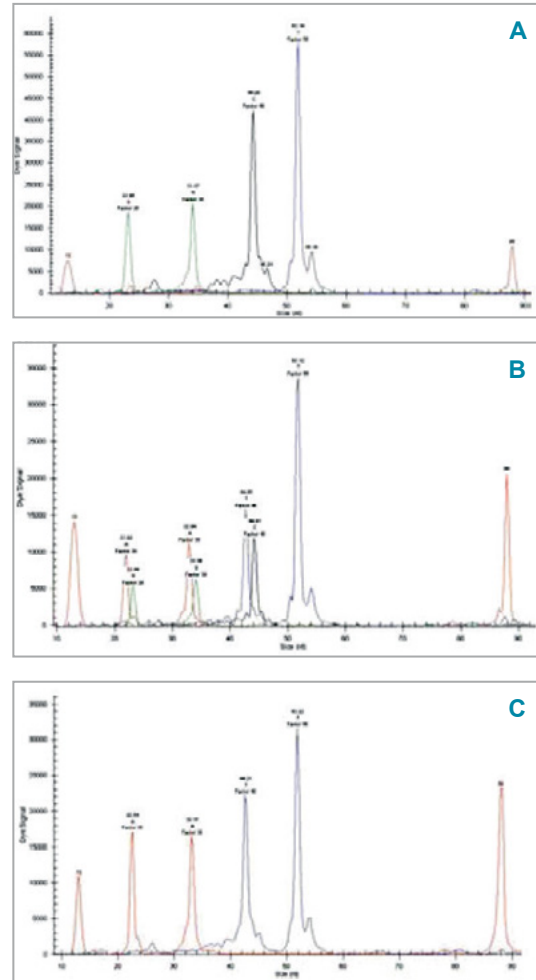
p5

[Download this technical note](#)

The software identifies and stores SNP locus information in SNP locus tags stored in the database. The SNP locus tags label the SNP fragment peaks with the correct size and the correct dye. SNP genotypes are summarized and reported in the fragment list through automated SNP locus tag assignment with a high degree of accuracy.

#### Factor V Mutation Scoring by GeXP Software

As shown in Figure 4, three primers (27 mer, 37 mer, and 47 mer) were synthesized to detect the allele identity at the factor V mutation site. A fourth primer (55 mer) was designed to serve as a positive control. True multiplex primer extension was performed using all four primers. The reactions added one nucleotide to each primer and resulted in products that were 28, 38, 48, or 56 nucleotides long, respectively. Genotypes of DNA template, wild-type, heterozygous, or homozygous for this G to A substitution locus, were scored singly or by multiplex primer extension assays and GeXP SNP analysis. Sample data were collected and analyzed using default SNP analysis parameters which were developed specifically for analyzing fragments smaller than 100 nucleotides. The apparent sizes for each extended fragment with actual size of 28, 38, 48, or 56 nucleotides were entered to create new locus tags. All sample data were then reanalyzed with four locus tags selected. As shown in Figure 5, the GeXP software was able to assign the correct allele ID and locus name to the correct peak. For heterozygous loci, the two dye-labeled fragments were labeled by two allele IDs with the same locus name. Over 500 SNP reactions were performed to score this factor V mutation. The GeXP was able to call the correct genotypes with a high degree of accuracy (almost 100%). The analyzed results display loci and alleles in the fragment list. Column selector and filter sets are available to edit the fragment list display (see A-1929A, "Strategies for Automating the Review of Data"). Genotypes for each result at each locus are summarized in the fragment list and could be directly printed, copied, or exported as a \*.CSV file into Microsoft Excel\* (Figure 6).



**Figure 4.** Correct allele calls for factor V mutation by GeXP software.  
A: Wild-type template;  
B: Heterozygous template at 28, 38, and 48 sites;  
C: Homozygous template at 28, 38, and 48 sites.

Utilize Column Selectors to select columns necessary for looking at SNP genotypes

Export the fragment list grid as \*.CSV file for SNP genotyping summary report

Standard SNP genotype summary includes:

- Result name
- Locus name
- Dye
- Allele ID
- Estimated size
- Peak height
- Allele matching quality

	A	B	C	D	E	F	G	H	I	J	K	L
1	RN	locus name	dye	allele ID	est frag size (nt)	pk height (riu)	allele match qual					
2	r7 g3 A04_02031515R3	FV 28	D3	G	25.79	2314	7.8					
3	r7 g4 B05_02031515R4	FV 28	D3	G	25.91	15890	6.03					
4	r7 g5 C06_02031515R4	FV 28	D3	G	25.48	19519	13.29					
5	r8 g2 A03_02031515R5	FV 28	D1	A	24.96	21546	14.23					
6	r8 g2 A03_02031515R5	FV 28	D3	G	25.86	11036	6.78					
7	r8 g4 C05_02031515R6	FV 28	D1	A	24.69	11130	12.81					
8	r8 g4 C05_02031515R6	FV 28	D3	G	25.99	11496	5.03					
9	r8 g5 D06_02031515R7	FV 28	D1	A	24.8	23056	10.93					
10	r8 g5 D06_02031515R7	FV 28	D3	G	25.72	13556	8.9					
11	r8 g2 B03_02031515R8	FV 28	D1	A	24.87	38286	12.4					
12	r8 g3 C04_02031515R9	FV 28	D1	A	24.69	14235	9.14					
13	r9 g5 E06_02031515Q0N	FV 28	D1	A	24.81	50391	11.23					
14	r7 g3 A04_02031515R3	FV 38	D3	G	35.8	2266	12.91					
15	r7 g4 B05_02031515R4	FV 38	D3	G	35.8	14923	12.96					
16	r7 g5 C06_02031515R4	FV 38	D3	G	35.8	17304	13					
17	r8 g2 A03_02031515R5	FV 38	D1	A	34.98	17764	8.94					
18	r8 g2 A03_02031515R5	FV 38	D3	G	35.81	8913	12.68					
19	r8 g4 C05_02031515R6	FV 38	D1	A	34.82	10807	6.52					
20	r8 g4 C05_02031515R6	FV 38	D3	G	35.82	12708	12.59					
21	r8 g5 D06_02031515R7	FV 38	D1	A	34.99	21578	9.04					
22	r8 g5 D06_02031515R7	FV 38	D3	G	35.81	13590	12.79					
23	r9 g2 B03_02031515R8	FV 38	D1	A	34.96	33183	8.66					
24	r9 g3 C04_02031515R9	FV 38	D1	A	34.9	11221	7.64					
25	r9 g5 E06_02031515Q0N	FV 38	D1	A	34.93	49654	8.14					
26	r7 g3 A04_02031515R3	FV 48	D2	C	45.93	4222	10.52					
27	r7 g4 B05_02031515R4	FV 48	D2	C	45.93	39078	10.53					
28	r7 g5 C06_02031515R4	FV 48	D2	C	45.9	47868	10.97					
29	r8 g2 A03_02031515R5	FV 48	D2	C	45.05	23600	12.02					
30	r8 g2 A03_02031515R5	FV 48	D4	T	45.78	11292	13.37					
31	r8 g4 C05_02031515R6	FV 48	D2	C	45.88	32948	11.36					
32	r8 g4 C05_02031515R6	FV 48	D4	T	45.78	15793	13.43					
33	r8 g5 D06_02031515R7	FV 48	D2	C	45.9	36814	11.1					
34	r8 g5 D06_02031515R7	FV 48	D4	T	45.77	18682	13.58					

**Figure 5.** SNP genotype summary report. A: Apply filter sets and edit columns to summary SNP genotypes in fragment list. B: Exported genotype report (\*.CSV) in Microsoft Excel.

Download this technical note

### Discussion

To achieve the best results with minimal chemistry optimization, several issues need to be considered:

- The migration of a short fragment during capillary separation is closely related to the length, sequence, and dye of this fragment. Both incorporation of dye and nucleotide composition greatly affect the mobility of the extension products. Often shorter fragments will appear to be nearly five bases longer than their actual size. We strongly recommend short primers be tested before being multiplexed to ensure enough adjacent space for data analysis. A difference of 3–5 nts between primer lengths is recommended.
- If signal over-ranging is observed during analysis, it may be necessary to dilute the samples with SLS prior to adding to the above mixture. Different scales of reactions with less enzyme, buffer, and dye terminators are also applicable to reduce signal and save reagents.
- Formation of hairpin or dimer structures among primers would lead to nonspecific signals after primer extension. To avoid this, we recommend performing a primer-only reaction prior to multiplexing.
- Incomplete removal of the free primer, dNTP, or ddNTP by SAP or Exo I digestion could result in nonspecific fragments. This problem can be avoided by using freshly mixed enzymes for each cleanup.

### Conclusions

The GeXP Genetic Analysis System, CEQ DNA Size Standard 80, and CEQ SNP-Primer Extension Kit enable medium- to high-throughput SNP validation and scoring.

- The SNP size standard, SNP dye mobility correction, and SNP locus tag were developed to ensure automated sizing and allele calls for SNP products.
- The GeXP Genetic Analysis System allows the analysis of multiple SNP loci in a single capillary.
- The GeXP Genetic Analysis System was able to validate and score human factor V mutation for venous thrombosis population with a high degree of accuracy

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. A-1928-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701, USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)







## Multiplex Mutation Genotyping for Human Diseases: Breast Cancer and Familial Mediterranean Fever

Zhiming Jiang,<sup>1</sup> Milhan Telatar,<sup>2</sup> Wayne Grody<sup>2</sup>

<sup>1</sup> Beckman Coulter, Inc.

<sup>2</sup> University of California, Los Angeles

### Introduction

Understanding human genetic variation is currently believed to reveal the cause of individual susceptibility to disease and the large variation observed in response to treatment. Single nucleotide polymorphisms (SNPs) are the subject of many novel technology development activities because of their high genome density and association with human diseases. The ideal SNP genotyping assay will be very accurate, inexpensive, easy to perform, and capable of high throughput.

Because all SNPs can be analyzed with high specificity at the same reaction conditions, primer extension is a promising reaction principle for multiplex high-throughput genotyping assays. In the primer extension reaction, a DNA polymerase is used specifically to extend a primer that anneals immediately adjacent to the polymorphic position with a single labeled dye terminator complementary to the nucleotide at the variant site. The reaction allows highly specific multiplexing detection of point mutations and SNPs.

The CEQ/GeXP Genetic Analysis System, DNA Size Standard 80, and SNP-Primer Extension Kit were developed to provide users an accurate, inexpensive, simple, and robust solution for multiplex SNP scoring and validation based on primer extension technology.<sup>1</sup> Our previous study<sup>1</sup> reported that single-base primer extension and the GeXP provided a novel way to screen factor V mutation in Venous Thrombosis disease population with high degrees of throughput, automation, and accuracy. Here we demonstrate the multiplexing analysis of two additional sets of human disease mutations, Ashkenazi Jewish BRCA1/2 mutations and MEFV mutations (Familial Mediterranean Fever), using the GeXP. The assays exhibit high accuracy and improved throughput through automation.

### Materials and Methods

#### Template Amplification

Templates were amplified from human genomic DNA by polymerase chain reaction (PCR<sup>\*</sup>). PCR conditions, thermal cycling conditions, and primer sequences are shown in Table 1, Figure 1, and Figure 3. The unincorporated PCR primers and dNTP were cleaned up by SAP and ExoI digestion<sup>1</sup> (Table 1). Prior to primer extension, templates were quantitated carefully by quantitative electrophoresis or laser-induced fluorescence using the Agilent<sup>\*</sup> 2100 Bioanalyzer (Cat. No. G2938A).

#### Primer Extension

Commercial software packages are available to design primers for multiplexing SNP analysis.<sup>2</sup> For successful SNP analysis, lengths of the adjacent primers for the same multiplexing reaction need to have a space of at least 5 nt. Regardless of the lengths, the Tms of all primers are within the range of 60° to 76° C—too great a difference in Tm would lead to large variations in the amount of extended products generated by different primers, therefore causing imbalanced signals in the multiplexing result. Adding a Poly (A) or Poly (T) tail to the 5'

end of primers longer than 25 nt is suggested to help minimize the Tm variations. All SNP primers targeting the BRCA1, BRCA2, or MEFV mutations are listed in Figure 1, Figure 3, and Table 2.

Primer extension was performed following the procedures presented previously.<sup>1</sup> Each primer was tested individually prior to multiplexing. As shown in Table 2, primer concentrations were adjusted to achieve the best signal balance in multiplexing reactions. The remaining dye terminators were hydrolyzed from the extended products by SAP digestion.<sup>1</sup> Extended products were then diluted in the sample loading solution (SLS) at a ratio of 1:10, added to size standard 80/SLS mixture<sup>1</sup> and finally loaded onto the GeXP.<sup>1</sup>

p1

[Download this technical note](#)

### Data Analysis

Data were analyzed using the SNP analysis feature of the GeXP software. SNP genotypes were summarized and reported in fragment list through automated SNP locus tag assignment.<sup>1</sup>

## Results and Discussion

### BRCA Mutations

The BRCA1 mutations (185delAG and 5382insC) and the BRCA2 mutation (6174delT) have been detected in a significant proportion of Ashkenazi Jewish women with early-onset breast cancer (Figure 1). In the Jewish Ashkenazi population, the overall frequency of these mutations is as high as 2.5%. Based on clinical and family history data in Ashkenazi individuals at risk, the results of BRCA molecular testing are appropriately anticipated in most cases. In order to screen the BRCA mutations using multiplexing primer extension, co-amplification of all three BRCA exons (two in BRCA1 gene and one in BRCA2 gene) was performed in a single PCR reaction (Table 1 and Figure 1).

SNP primers that are 23 nt, 33nt, or 43 nt long, respectively, were designed to hybridize one base pair adjacent to the corresponding mutation sites (Table 2 and Figure 1). Primer extension was performed, and all three primers were able to detect the wild type or the mutant allele of each mutation locus with a high degree of accuracy. To achieve the best multiplexing results, primers were tested individually prior to multiplexing. Primer concentrations were adjusted to balance signal levels generated by the different primers in the same multiplexing result (Table 2). The extended products were cleaned up prior to separation on the GeXP.<sup>1</sup> SNP data were analyzed using the fragment analysis module of the GeXP.<sup>1</sup> Three locus tags were created according to the apparent sizes of SNP products (Table 2). For heterozygous loci, two allele IDs with the same locus name were labeled. Figure 2 exhibits typical examples of the analyzed BRCA data from a wild-type person and a person heterozygous for the 5382InsC locus. Overall, multiplexing primer extension and the GeXP system were able to detect BRCA mutations with a high degree of accuracy and minimal optimization in both chemistry and software.

Gene	PCR Reaction	Cycling Condition	Template Cleanup
BRCA1 and BRCA2 (co-amplification of three exons)	DNA	100 ng	94° C 1 min
	BRL buffer (10X)	3 µL	
	MgCL2 (50 mM)	3.9 µL	30 cycles:
	dNTP (10 mM)	0.3 µL	94° C 30 sec
	Each primer (100 µM)	0.1 µL	58° C 30 sec
	Platinum Taq	0.4 µL	72° C 3 min
	H <sub>2</sub> O	q.s.	
	Total Volume	30 µL	
MEFV (co-amplification of three exons)	DNA	100 ng	94° C 1 min
	BRL buffer (10X)	3 µL	
	MgCL2 (50 mM)	3 µL	30 cycles:
	Betaine (5 mM)	12.6 µL	94° C 30 sec
	dNTP (10 mM)	0.4 µL	60° C 30 sec
	Each primer (100 µM)	0.1 µL	72° C 50 sec
	Platinum Taq	0.4 µL	
	H <sub>2</sub> O	q.s.	72° C 3 min
Total Volume	30 µL		
			To 30 µL PCR product:
			10 Unit SAP
			5 Unit ExoI
			37° C 1 hour
			75° C 15 min

**Table 1.** PCR Conditions for Template Amplification.

Disease	Gene	Primer Name	Primer Length (nt)	Mutation Identity	Amount Used for Each Multiplexing SNP Reaction	T <sub>m</sub> (°C)	Apparent Size (nt)
Breast Cancer	BRCA1	P33 185delAG	3	185delAG	0.8 pmol	61.5	19
	BRCA1	P23 5382InsC	23	5382InsC	2.0 pmol	68.8	31
	BRCA2	P43 6174delT	43	6174delT	2.0 pmol	61.5	41
Familial Mediterranean Fever	MEFV	P23 M680I	23	M680I	1.0 pmol	67.5	20
	MEFV	P55 M694V	55	M694V	1.0 pmol	65.0	57
	MEFV	P37 M694I	37	M694I	0.5 pmol	66.7	37
	MEFV	P47 V726A	47	V726A	0.5 pmol	69.9	47
	MEFV	P28 E148Q	28	E148Q	1.0 pmol	76.0	24

**Table 2.** SNP Primers Designed for Each Disease Mutation Site.

p2

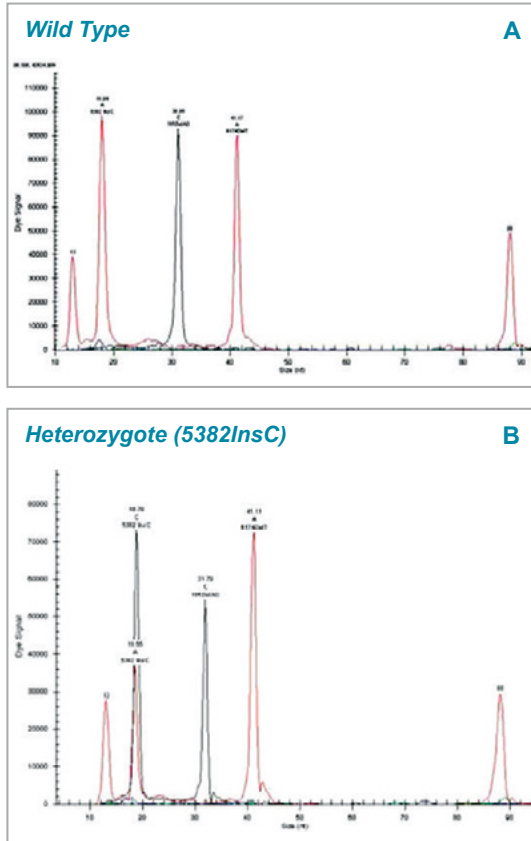
[Download this technical note](#)

**1A BRCA1: 185delAG mutation**  
 PCR primer (Forward)  
**CTTCGCGTTG AAGAAGTACA** AAATGTCATT AATGCTATGC  
 Deletion  
 AGAAAATCTT **AG** AGTGTCCCATCTGGTAAGTCAGCACAAG  
 ←  
 AGTGTATTAA TTTGGGATTCTATGATTAT CTCCTATGCA  
 — A: P33 185delAG  
 AATGAACAGAATTGACCTT **ACATACTAGGGAAGAAAAGAC A**  
 PCF primer (Reverse)

**1B BRCA1: 5382InsC mutation**  
 PCR primer (Forward)  
 TGGAAGAAACCACCAAGGTC CAAAGCGAGCAAGAGAAT  
 Insertion C P23 5382INSCT: —  
 CCC **X** AGGACAGAA AGGTAAAGCTCCCTCCCTCAAGTTGACAAA  
 →  
 AATCTCACCC **CACCACTCTG TATTCCA** **CTC CCC**  
 PCF primer (Reverse)

**1C BRCA2: 6174delT**  
 PCR primer (Forward)  
**GGGAAGCTTC ATAAGTCAGT** CTCATCTGCAA TACTTGTG  
 Deletion  
 GGATTTTTAG CACAGCAAG **T** GGAAAATCTG TCCAGGTATC  
 ←  
 AGATGCTTCA TTACAAAACG CAAGACAAGT GTTTTCTGAA  
 — T: P43 6174delT  
 ATA **GAAGATA GTACCAAGCA** **AGTC**  
 PCF primer (Reverse)

**Figure 1.** Sequences of templates, PCR primers, and SNP primers for the BRCA mutations. (The PCR primers are bolded and underlined. Each SNP primer name contains the corresponding mutation identity and the length of the primer.) 185delAG locus (SNP primer in green) 5382insC locus (SNP primer in orange) 6174delT locus (SNP primer in blue).



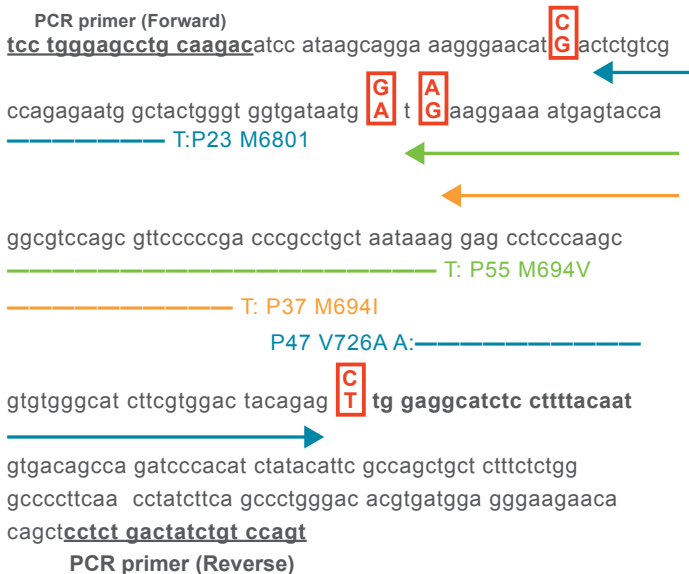
**Figure 2.** Typical multiplexing genotyping results for BRCA mutations (the SNP locus tag was named after the mutation identity). Red peaks (13 and 88) are size standards.

### MEFV Mutations

Familial Mediterranean Fever (FMF) is an autosomal recessive disorder characterized by recurring attacks of fever and serositis. Five sequence alterations (M694V, V726A, M680I, M694I, and E148Q) in the MEFV gene account for the majority of FMF cases. The MEFV gene, located on the short arm of chromosome 16, codes an anti-inflammatory protein, called marenostriine or pyrin. As shown in Figure 3, four out of the five mutations (M694V, V726A, M680I, and M694I) are located on exon10 of the MEFV gene, whereas the fifth mutation (E148Q) is on exon 2 of the MEFV gene. Since exon 2 contains a high percentage of GC nucleotides, betaine (a final concentration of 2.1 mM) was added to help co-amplification of both exon 10 and exon 2 templates in one PCR reaction (Table 1). PCR products were cleaned and quantitated prior to primer extension as described in "Materials and Methods." Five SNP primers (23, 55, 37, 47, 28 nt long, respectively) were designed to score five MEFV mutations correspondingly (Table 2 and Figure 3). Within the exon10 template, two mutations (M694V and M694I) are physically only one base pair apart. Two overlapping primers (55 nt and 37 nt) were proven to be able to score these two mutations respectively (Table 2 and Figures 3 and 4). A relatively short primer (28 nt) was used to screen the E148Q mutation in the GC-rich template. As shown in Figure 4, multiplexing SNP analysis using primer extension and the GeXP was capable of genotyping individuals that are wild type, heterozygous, or homozygous at all five FMF loci. Primer extension templates from each individual were co-amplified in one single PCR reaction. Five loci were screened by one multiplexing primer extension.

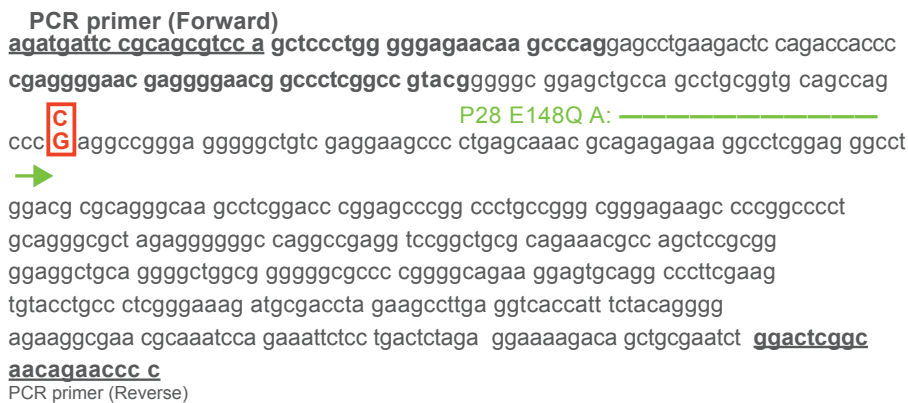
3A

MEFV Exon10: M6801, M694V, M694I, and V726A



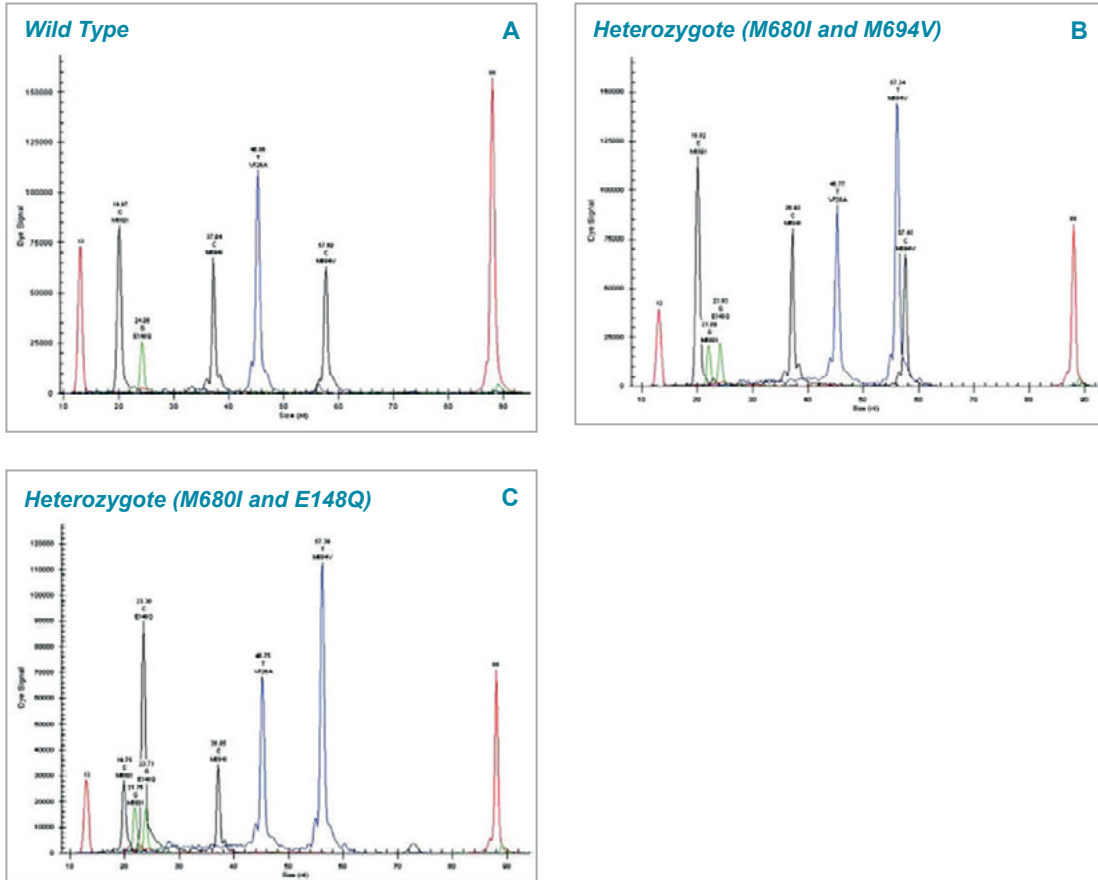
3B

MEFV Exon 2: E148Q



**Figure 3.** Sequences of templates, PCR primers, and SNP primers for the MEFV mutations. (The PCR primers are bolded and underlined. Each SNP primer name contains the corresponding mutation identity and the length of the primer.) Four mutations in Exon10: M6801 (primer in blue), M694V (primer in green), M694I (primer in orange), and V726A (primer in blue); One mutation in Exon2: E148Q (primer in green).





**Figure 4.** Typical multiplexing genotyping results for BRCA mutations (the SNP locus tag was named after the mutation identity). Red peaks (13 and 88) are size standards.

[Download this technical note](#)

### Conclusion

The GeXP, DNA Size Standard 80, and SNP-Primer Extension Kit enable validation and scoring of multiple sets of disease mutations (BRCA and MEFV) with high degrees of accuracy, throughput, and automation. The GeXP was able to handle problematic loci with minimal requirements in assay optimization. Both template amplification and primer extension assay are amenable to future automation with no requirement for equipments other than common lab centrifuges and vacuum manifolds.

### References

1. Jiang, Zhiming, Telatar, Milhan, Grody, Wayne. Multiplex SNP Analysis: Screening Factor V R506Q (Leiden) Mutations. *Application Information Bulletin A-1928A*, Beckman Coulter, Inc.
2. Rychlik, Wojciech. *Oligo Primer Analysis Software User Manual*. Molecular Biology Insights, Inc.

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. A-1942A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701, USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)



# Gene expression profiling

## Quantitative Multiplexed Gene Expression: An Analysis of Sensitivity, Accuracy and Precision

Yong Wu, Mervin R. Gutierrez, Kathryn E. Sciabica, Jane Luo  
SCIEX

### Introduction

Quantitative gene expression analysis is considered the gold standard by which all gene expression analysis tools are measured. The GenomeLab™ GeXP Genetic Analysis System from SCIEX offers multiplexed, quantitative gene expression analysis capable of examining up to 30 genes in a single reaction from as little as 5 ng total RNA. Here, we demonstrate that the GeXP produces gene expression data with superb linearity and is sensitive enough to precisely detect even small changes in gene expression. A highly linear correlation between the amount of RNA and gene expression quantitation value was generated for each gene in a multiplex with an average correlation coefficient ( $R^2$ ) well above 0.99. In addition, we verified that ten consecutive 0.5-fold increases in RNA concentration were accurately quantified by the GeXP for all 24 genes in the multiplexed assay. The capacity of GeXP to deliver multiplexed, sensitive and precise gene expression analysis opens a new door for scientists to explore subtle yet biologically meaningful changes in an effective and efficient manner.

Quantitative analysis of multiplexed gene expression in a single reaction, from a limited amount of total RNA, is of great interest to research scientists. Conventional techniques utilize either microarrays, which detect the expression of a high number of genes per reaction at high cost and low sample throughput, or real-time PCR\* assays, which detect the expression of a few genes at moderate throughput. More often researchers have shifted their focus to 20 or more genes during biomarker validation, signal transduction or metabolic pathway studies. Following this shift, SCIEX has recently developed the GenomeLab GeXP (Gene eXpression Profiler) Genetic Analysis System that can detect up to 30 genes in a single reaction

containing 5 ng to 500 ng of total RNA. GeXP uses a patented universal priming strategy to overcome potential bias in amplified targets that are typically associated with other types of multiplexed assays. Since its inception, GeXP has demonstrated correlative results to real-time PCR and microarray data. <sup>(1-4 and unpublished data)</sup>

Equipped with eight capillaries and a two-plate format, the GeXP system is capable of analyzing 240 samples per 24- hour period. This produces 7,200 gene expression results for a 30-gene multiplexed assay. Researchers not only benefit from reduced labor and reagent expenditures, but also significant time savings due to expedited data collection with the GeXP system. Further, by providing more gene expression results per sample than real-time PCR, GeXP diminishes the experimental constraint imposed by limited quantities of RNA, such as those from formalin-fixed paraffin-embedded (FFPE) samples or tissue biopsies. The capabilities of GeXP has been increasingly recognized by academic and other research communities ranging from plant sciences and viral infection studies to cancer research.<sup>1-4</sup>

In this study, we demonstrate that GeXP has precise linearity in gene expression quantitation for each gene in a multiplex. The gene expression quantitation (GEQ) value, calculated from a standard curve, displayed highly linear correlation to the amount of RNA used. The average correlation coefficient ( $R^2$ ) achieved was greater than 0.99. We further confirm that the GeXP delivers unique sensitivity in detecting subtle changes in RNA levels. The overall results indicate that the GeXP system is not only powerful enough to effectively deliver thousands of gene expression results per day, but also sensitive enough to detect small changes of gene expression levels in a very precise manner.

## Materials and Methods

**Reagents:** GenomeLab GeXP Start Kit and GenomeLab GeXP Human Breast CancerPlex Kit (Beckman Coulter, Inc.), ThermoStart\*\* Taq DNA Polymerase (Thermo Scientific)

**Reverse transcription, multiplex PCR and PCR fragment separation:** cDNA synthesis from the reverse transcription of total RNA, multiplex PCR and PCR fragment separation on the GeXP system were performed following the standard GeXP Chemistry Protocol (P/N A29143) and GeXP Human Breast CancerPlex Kit Insert (P/N A32698).

### Optimization of the amount of KANr RNA in GeXP reactions:

The data analysis method presented in this study utilized the signal level of a gene relative to the internal control, KANr gene. The original kanamycin resistance (KANr) RNA (5ng/uL) from the GenomeLab GeXP Start Kit was diluted 2-fold over 11 serial dilutions with RNA Storage Solution (Ambion). The final amount of KANr RNA in each reaction was in a range from 25 pg to 25 ng, and was added into a GeXP multiplex reverse transcription reaction containing 20 ng of Human Breast CancerPlex (HuBC) control RNA from the Human Breast CancerPlex Kit. The amount of KANr RNA that generated a peak height signal intensity equivalent to the median signal levels in the multiplex was selected. This optimal amount of KANr RNA (0.1 ng) was used in the reactions to establish the standard curves.

**Establishment of standard curves:** GeXP multiplex reactions were carried out using decreasing amounts (from 512 ng to 1 ng in 2-fold dilutions) of HuBC control RNA in the presence of the optimized amount of KANr RNA, 0.1 ng. For each amount of control RNA, six technical replicates were prepared for multiplex reverse transcription. Each multiplex reverse transcription reaction was independently used for a multiplex PCR reaction. Each multiplex PCR reaction was separated on a different capillary channel of a capillary array.

**Determination of GeXP sensitivity in quantitation:** A series of ten 0.5-fold increase of HuBC control RNA, from 6.5 ng to 250 ng were used in reactions containing 0.1 ng of KANr RNA. For each amount of control RNA, four technical replicates were prepared for multiplex reverse transcription. Each multiplex reverse transcription reaction was independently used for a multiplex PCR reaction. Each multiplex PCR reaction was separated on a different capillary channel of a capillary array.

**GeXP fragment analysis:** Fragments separated on the GeXP system were analyzed by a set of modified GeXP fragment analysis parameters (slope threshold = 1, peak height threshold

= 0) in order to call all possible peaks including the smallest ones. Multiplex-specific fragments were selected by applying exclusion filters and were exported to eXpress Analysis software, where they were normalized against the KANr gene. Normalized data was exported to Microsoft Excel\*\* for further analysis.

**Data analysis:** An XY (scatter) chart was used to plot the relationship between the relative signal level normalized against KANr (X-axis) and the amount of HuBC control RNA (Y-axis). A best fit regression line with a third-order polynomial was applied. The regression line was considered a standard curve and its equation was used to calculate the GEQ value from relative signal level normalized against KANr obtained from the eXpress Analysis report. The relative accuracy (RA) of a GEQ value is defined as:  $RA = (RNA\ amount - Absolute\ (RNA\ amount - GEQ\ value)) / RNA\ amount \times 100\%$ .

## Results

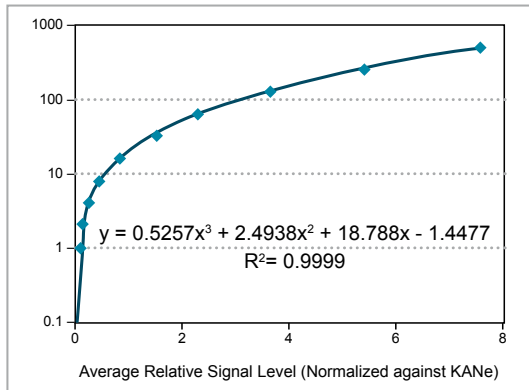
Establishment of the standard curve for a multiplex Like many other quantitation methods, the quantitative analysis developed in this study uses a standard curve to calculate the GEQ value corresponding to a relative signal level. An overall standard curve was established to evaluate how precisely the average signal of a multiplex correlates to the amount of RNA. To establish the standard curve for Human Breast CancerPlex, the fragment data from GeXP reactions performed with 2-fold serial dilutions of HuBC control RNA, ranging from 512 ng to 1 ng, were analyzed by eXpress Analysis software. Relative signal level of each gene was obtained by normalizing its peak area against that of the KANr gene. These relative signal levels were exported to Microsoft Excel from eXpress Profiler in the Profile (by gene) format. The average of the relative signal level of all 24 genes for each concentration of control RNA was then calculated.

A scatter chart was used to plot the relationship between the average of relative signal level of each multiplex reaction (x-axis) and the amount of HuBC control RNA (y-axis). A best fit regression curve with a third-order polynomial was applied (Figure 1). A coefficient of determination ( $R^2$ ) of 0.9999 was obtained between the average relative signal level and the amount of HuBC control RNA. In regression analysis, the  $R^2$  is a statistical measure of how well the regression line approximates the real data points. It represents the fraction of variability in the y-values that can be explained by the variability in the x-values. An  $R^2$  of 1.0 indicates that the regression line perfectly fits the data. Therefore, a regression curve with high  $R^2$  value, generally above 0.99, in GeXP data analysis can be considered a very reliable standard curve. The equation of the standard curve

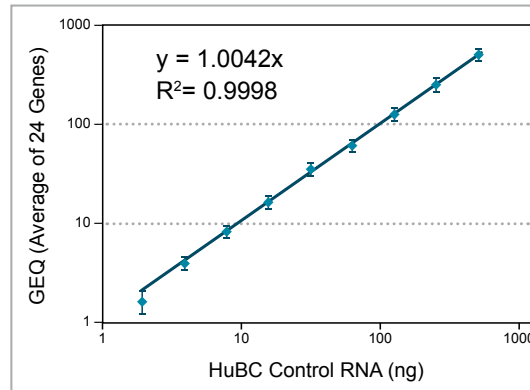
p2

[Download this technical note](#)





**Figure 1.** The correlation between the average of relative signal level for six technical replicates of all 24 genes in the Human Breast CancerPlex and the amount of HuBC control RNA was near perfect in a third-order polynomial model. The fitting equation and  $R^2$  are shown on the chart. The GEQ value (y) is calculated using the equation from this standard curve with relative signal level (x) for a particular RNA concentration.



**Figure 2.** The linear correlation between the amount of HuBC control RNA and the GEQ value. The relationship between the amount of HuBC control RNA and the GEQ value from the average of relative signal level of all 24 genes in Human Breast CancerPlex was plotted. A linear regression model was added to data points from 2 ng to 512 ng total input RNA. Error bars represent standard deviation of six technical replicates. The correlation coefficient ( $R^2$ ) is displayed on the chart.

(regression curve) is used to calculate the gene expression quantitation (GEQ) value (y) from a relative signal level (x) obtained from the eXpress Analysis report. Therefore, a GEQ value is a signal level quantitation derived from a standard curve.

The standard curve equation (Figure 1) for Human Breast CancerPlex was used to calculate the GEQ value (y) from the relative signal level (x) for each technical replicate. When the average GEQ value from six technical replicates was plotted against the amount of HuBC control RNA, a superb linear correlation was achieved. The correlation coefficient ( $R^2 = 0.9998$ ) between the amount of HuBC control RNA and GEQ value indicates a highly precise linear relationship between amount of RNA and the corresponding GEQ value calculated from the standard curve. (Figure 2).

To evaluate the accuracy of this data analysis method, the relative accuracy (RA) for each GEQ value was calculated using the equation described in the Materials and Methods. The average RA is 95.0% for the GEQ value corresponding to the amount of HuBC control RNA from 2 ng to 512 ng (Table 1). A smaller average RA value was observed when the amount of RNA was less than 4 ng, primarily due to lack of signal for some

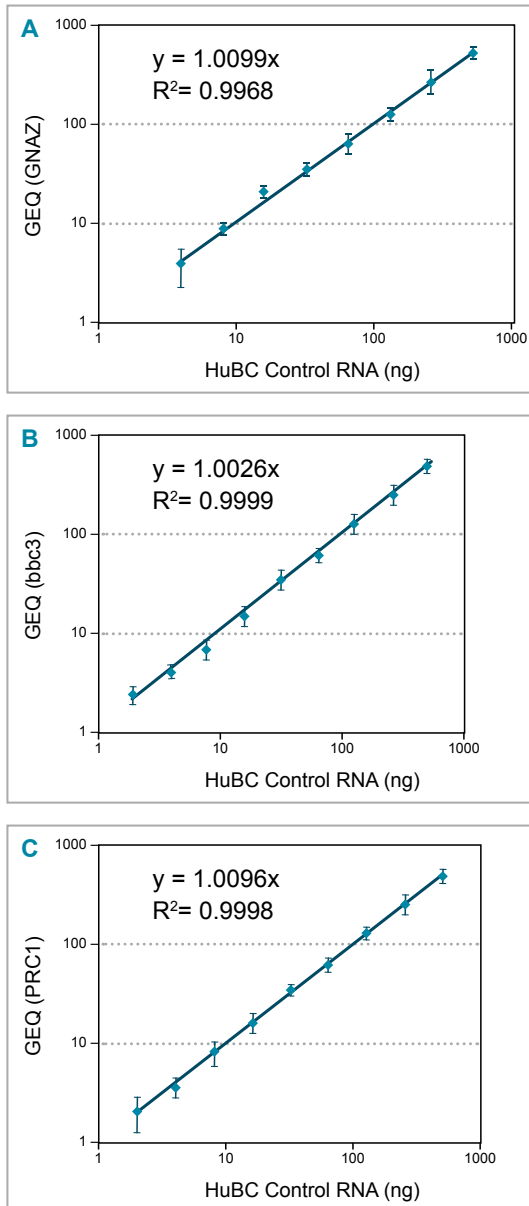
of the low expressers in the Human Breast CancerPlex. The average RA increases to 97% if only data points from 4 ng to 512 ng of RNA are considered.

This approach also provides a high quality fitting model for each individual gene in the Human Breast CancerPlex. By plotting the relative signal level of an individual gene against amount of HuBC control RNA, a trend line with a specific equation is obtained for each gene in the multiplex. This equation was then used to calculate the GEQ value for each gene. To briefly demonstrate how well this analysis works on individual genes, we selected three genes from the multiplex: a low expresser, a median expresser and a high expresser.

Gene GNAZ (NM\_002073), a low expresser, shows a very strong correlation between the amount of HuBC control RNA and the GEQ value from 8 ng to 512 ng of HuBC control RNA (Figure 3a). A near perfect correlation coefficient ( $R^2$ ) was also observed between amount of HuBC control RNA and the GEQ value within RNA range from 2 ng to 512 ng for a median expresser *bbc3* (U82987, Figure 3b) and a high expresser PRC1 (NM\_003981, Figure 3c).

HuBC control RNA (ng)	2	4	8	16	32	64	128	256	512
GEQ value	1.6	4.1	8.2	16.5	35.5	61.4	126.5	257.6	511.6
RA (%)	79.4	98.2	97.7	97.0	89.0	95.9	98.8	99.4	99.9
Average RA %	95.0								

**Table 1.** Relative Accuracy for the GEQ value.



**Figure 3.** The linear correlation between the amount of HuBC control RNA and the GEQ value for gene GNAZ (a), bbc3 (b) and PCR1 (c). The relationship between the amount of RNA and the GEQ value for each gene was plotted with a linear trendline. Error bars represent one standard deviation from the mean for six replicates. The  $R^2$  of each gene is displayed on the respective chart.

### Determination of GeXP sensitivity in quantitation

To determine if GeXP is capable of detecting small changes in levels of gene expression, a 10-point dilution series varying the amount of HuBC control RNA in 0.5-fold increments, from 6.5 ng to 250 ng, was tested. A high correlation coefficient ( $R^2 = 0.9978$ ) between the amount of RNA and the GEQ value was obtained (Figure 4). The %CV of this experiment, which reflects total variation in the GeXP process (reverse transcription, PCR and capillary electrophoresis), ranges from 6.6 to 16.5. The average RA of this assay is higher than 96% (Table 2). This result indicates that GeXP can confidently detect gene expression changes as small as 0.5-fold.

Each of the 24 genes in Human Breast CancerPlex was analyzed with this method and the  $R^2$  for each gene was obtained (Table 3). For all 24 genes, the average and median values for the  $R^2$  were 0.9955 and 0.9976, respectively. These results demonstrate that GeXP delivers highly sensitive and precise quantitative gene expression analysis.

### Discussion

Sensitivity, precision and accuracy are key aspects of quantitative gene expression analysis. In this study, our results demonstrate that the GenomeLab GeXP possesses high standards for all of these qualities in a multiplexed approach.

The sensitivity to detect minute yet biologically important changes in gene expression has been largely desired by scientists to study the impact of subtle gene regulation, particularly in research areas of cancer, neurosciences and obesity.<sup>5-8</sup> Currently, small changes in gene expression cannot be accurately measured by conventional methods. Therefore, the significance of these subtle changes during tumor progression and neurodegeneration is poorly understood. The sensitivity of GeXP brings gene expression analysis into exciting new territory where subtle yet biologically significant changes are quantified, leading to the discovery of finely tuned mechanisms of gene expression regulation.

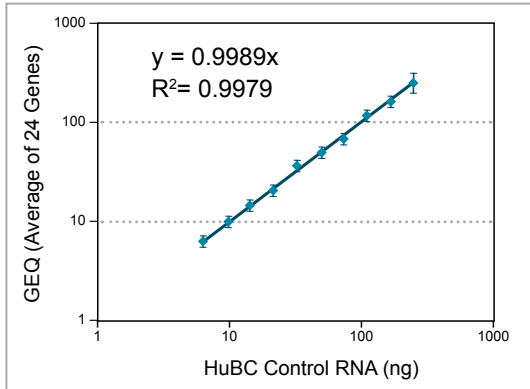
Our results demonstrate that GeXP is more sensitive in detecting minute changes in levels of gene expression than real-time PCR. This is likely due to the much higher signal to noise ratio of GeXP. Both GeXP and real-time PCR have their lowest signal to noise ratio at the beginning of PCR and both generate increasingly stronger signal as the cycle number progresses. Real-time PCR generally analyzes data points at a signal level ten times higher than the standard deviation of background

p4

[Download this technical note](#)

HuBC control RNA (ng)	6.5	9.8	14.6	21.9	32.9	49.4	74	111	167	250
GEQ value	6.5	9.8	15.0	20.4	36.2	49.8	67.8	118.3	162.7	250.6
RA (%)	99.8	99.9	97.2	93.3	89.9	99.3	91.6	93.4	97.4	99.8
Average RA %	96.2									

**Table 2.** Relative Accuracy (RA) of GeXP in detecting 0.5-fold change in amount of HuBC control RNA.



**Figure 4.** The linear correlation between the amount of HuBC control RNA and the GEQ value on detecting 0.5-fold change of target mRNAs. Error bars represent one standard deviation from the mean for four technical replicates. The correlation coefficient ( $R^2$ ) is shown on the chart.

noise, known as the threshold cycle ( $C_t$ ). In contrast, GeXP collects data points at a signal level a thousand to million times higher than background noise. Data collection at the lower signal to noise ratio may contribute to higher standard deviation for real-time PCR data. The user's guide for performing real-time PCR<sup>9</sup> specifies that the standard deviation of  $C_t$  value among technical replicates should not be greater than 0.3. A standard deviation of  $C_t$  value close to 0.3 reflects approximately 70% variation among technical replicates. In order for real-time PCR to be able to detect a 0.5-fold increase in mRNA level, the standard deviation of  $C_t$  value must be smaller than 0.1, which is very challenging under most circumstances.

A linear correlation between the GEQ value and the amount of RNA for each gene was achieved in GeXP multiplexed reaction with the average  $R^2$  well above 0.99. Combined with the high signal to noise ratio, this linearity enables GeXP to obtain a very high RA value of approximately 95%. In an evaluation study, real-time PCR demonstrated an RA value of 9.4, which is equivalent to an RA of 90.6% using the calculation method in this study.<sup>10</sup> The higher RA of GeXP over real-time PCR confirms that GeXP is not only geared for cost and time savings, but also delivers higher accuracy than conventional quantitative

Gene Symbol	Accession#	$R^2$
PPIA	BC000689	0.9984
GAPDH	NM_002046	0.9992
IGFBP5	L27560	0.9992
CP	NM_000096	0.9994
ESR1	NM_000125	0.9996
EXT1	NM_000127	0.9942
KRT18	NM_000224	0.9965
OXCT1	NM_000436	0.9985
DCK	NM_000788	0.9927
ACTB	NM_001101	0.9964
AP2B1	NM_001282	0.9945
GNAZ	NM_002073	0.9869
MYBL2	NM_002466	0.999
RFC4	NM_002916	0.9989
TGFB3	NM_003239	0.9993
CDC42BPA	NM_003607	0.9862
ALDH4A1	NM_003748	0.9977
WISP1	NM_003882	0.9811
PRC1	NM_003981	0.9963
KNTC2	NM_006101	0.9926
RAB6B	NM_016577	0.9913
EGLN1	NM_022051	0.9988
bbc3	U82987	0.9976
HSC4A2 IV2	X05610	0.9976
Average		0.9955
Median		0.9976

**Table 3.** For each gene in Human Breast CancerPlex in detecting 0.5-fold changes

methods. Another advantage of GeXP over one gene at a time assays is its accuracy in normalization. Because reference genes are analyzed in the same reaction with genes of interest, normalization in GeXP is not affected by pipette variation or well-to-well variation during the PCR process. Further, the ability of GeXP to include multiple housekeeping genes in one multiplex provides users the flexibility to choose a different housekeeping gene or use multiple housekeeping genes for a particular study.

Lower than average RA was found for the GEQ values corresponding to the amount of total RNA less than 4 ng for some low expresser genes in the multiplex. This can be attributed to the fact that when the amount of total RNA is low, there are too few copies of low expresser mRNA to generate reproducible signals. We recommend using a minimum of 5 ng of RNA for the multiplex reaction. This minimum amount remains much lower than the one required by real-time PCR and microarray assays. For median and high expressers, 2 ng of total RNA is usually enough to obtain reproducible results.

In summary, GeXP demonstrates superb linearity and accuracy in analyzing gene expression. Its ability to detect small but biologically significant changes distinguishes GenomeLab GeXP from other currently existing technologies. With the capacity of analyzing up to 30 genes in a single reaction, GeXP is equipped to deliver accurate and precise information in the most effective and efficient manner.

### Acknowledgement

We greatly appreciate Jeff Chapman, Samira Kaissi, Sam Dougaparsad, Amber Volk and many others for reviewing the manuscript.

### References

1. Chen QR, Vansant G, Oades K, Pickering M, Wei J, Song YK, Monforte J, and Khan J. 2007 Diagnosis of the small round blue cell tumors using multiplex polymerase chain reaction. *Journal of Molecular Diagnostics* 9: 80-88.
2. Boyer JD, Robinson TM, Kutzler MA, Vansant G, Hokey DA, Kumar S, et. al. 2007 Protection against simian/human immunodeficiency virus (SHIV) 89.6P in macaques after coimmunization with SHIV antigen and IL-15 plasmid. *Proceedings of the National Academy of Sciences* 104: 18648-18653.
3. Hayashi E, Aoyama N, Wu Y, Chi HC, Boyer SK, and Still DW. 2007 Multiplexed, quantitative gene expression analysis for lettuce seed germination on GenomeLab™ GeXP Genetic analysis System. *Application Information*, A-10295A, Beckman Coulter Inc. <http://www.beckmancoulter.com/Literature/BioResearch/A-10295A.pdf>
4. Kim B, Masuta C, Matsuura H, Takahashi H, and Inukai T. 2008 Veinal necrosis induced by Turnip mosaic virus infection in Arabidopsis is a form of defense response accompanying HR-like cell death. *Molecular Plant-Microbe Interactions* 21: 260-268
5. Koza RA, Nikonova L, Hogan J, Rim JS, Mendoza T, Faulk C, et. al. 2006 Changes in gene expression fore shadow dietinduced obesity in genetically identical mice. *PLoS Genetics* 2: 0769-0780
6. Yan H, Dobbie Z, Gruber SB, Markowitz S, Romans K, Giardiello FM, et. al. 2002 Small changes in expression affect predisposition to tumorigenesis. *Nature Genetics* 30: 25-26
7. Kaisera S, Foltza LA, Georgea CA, Kirkwoodb SC, Bemisc KG, Lind X, et. al. 2004 Phencyclidine-induced changes in rat cortical gene expression identified by microarray analysis: implications for schizophrenia. *Neurobiology of Disease* 16: 220-235
8. Bennett CN, Ross SE, Longo KA, Bajnok L, Hemati N, Johnson KW, et al. 2002 Regulation of Wnt signaling during adipogenesis. *Journal of Biological Chemistry* 277: 30998–31004.
9. Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR. Applied Biosystems.
10. Canales RD, Luo Y, Willey JC, Austermler B, Barbacioru CC, Boysen C, et. al. 2006 Evaluation of DNA microarray results with quantitative gene expression platforms. *Nature Biotechnology* 24: 1115-1122.

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

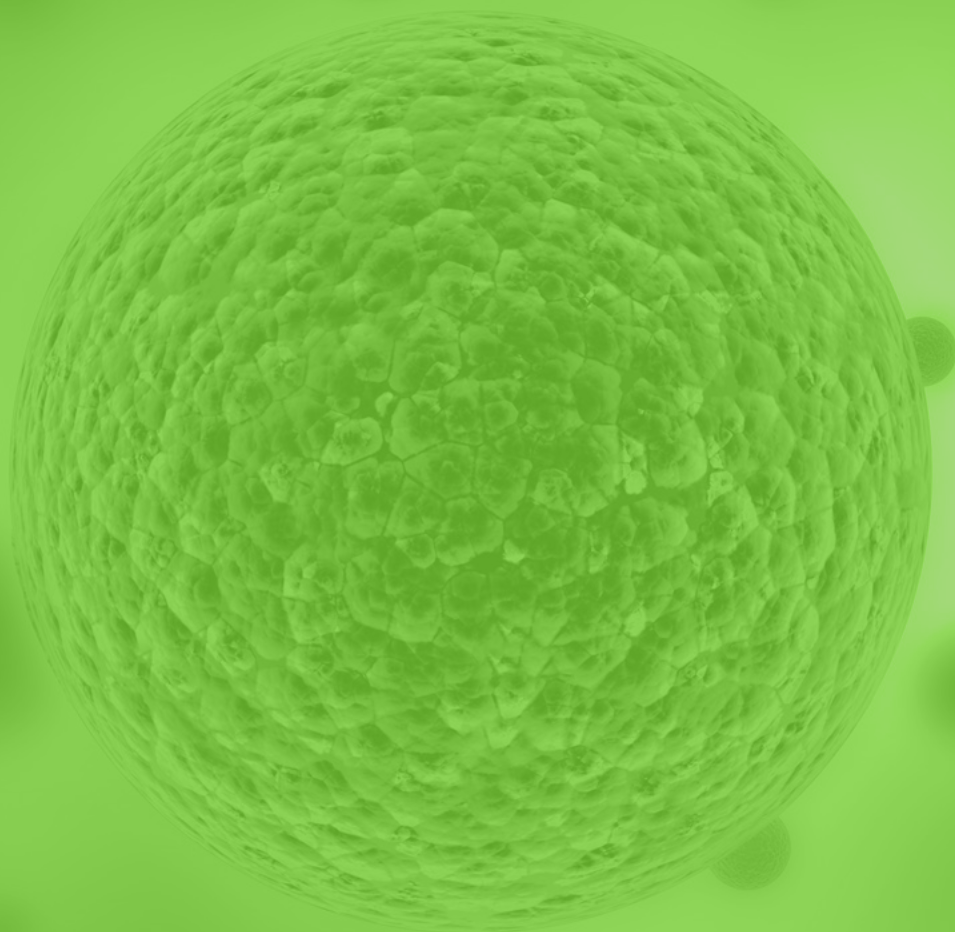
AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. T-11555A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701, USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

Download this technical note



## Monitoring Multiple Pluripotency Biomarkers After Delivery of Reprogramming Factors to Human Somatic Cells

Jane Pham,<sup>1</sup> Carl Javier,<sup>1</sup> Kyle Howerton,<sup>1</sup> Joel Marh,<sup>1</sup> Jessie Kinjo,<sup>1</sup> Jadelind Wong,<sup>1</sup> Michael Pascual,<sup>1</sup> Fariborz Izadyar,<sup>1</sup> Paul Ludtke,<sup>1</sup> Kathryn Sciabica,<sup>2</sup> Christian Kannemeier<sup>1</sup>  
*1 PrimeGen Biotech LLCa, 213 Technology Drive, Irvine, CA 92618*  
*2 Beckman Coulter, Inc.*

### Summary

Trans-acting factors present in the mammalian oocyte are capable of reprogramming somatic cell nuclei to an undifferentiated state and produce clonal embryonic stem cells (ESCs) in a process called somatic cell nuclear transfer. Recent studies demonstrate that a basic set of transcription factors (Oct4, Sox2, cMyc and Klf4) are sufficient to reprogram mouse somatic cells to pluripotent stem cells that exhibit the essential characteristics of ES cells. It is anticipated that such induced pluripotent stem (iPS) cells will be useful for development of disease models, drug studies and eventually patient-specific tissue replacement. This investigation demonstrates that expression of these four factors, plus the additional transcription factor Nanog, from recombinant lentivirus, effectively reprogrammed human somatic cells such that they displayed a pluripotent phenotype and expressed multiple pluripotency genes. It was further demonstrated that the genes for these transcription factors were transiently transduced into human somatic cells with nonviral, particle vectors and transiently induced expression of pluripotency markers.

### Introduction

ESCs are unique in their ability to grow indefinitely while maintaining pluripotency. Their ability to differentiate into virtually any cell type makes them prime candidates for regenerative medicine. However, ethical concerns surround the origins of human ESCs and practical concerns relate to the potential for immune rejection. It has been reported that pluripotent stem cells can be induced from mouse and human adult fibroblasts by introducing four transcription factors, Oct4, Sox2, cMyc, and Klf4, under ESC culture conditions using retroviral vectors.<sup>1,2</sup> Such iPS cells have great potential to support regenerative and developmental research in conjunction with human ESCs.<sup>3,4</sup> It is thought that both could ultimately lead to personalized cell replacement therapies.<sup>5</sup> However, the use of retroviral vectors in a clinical environment is likely to be problematic due to the

increased risk of tumorigenicity with genome-integrating viruses. Therefore, non-viral delivery mechanisms are highly desirable. This investigation sought to determine whether various types of human somatic cells could effectively be reprogrammed in the presence of five transcription factor genes (Oct4, Sox2, cMyc, Klf4 and Nanog) delivered either by lentivirus vectors or non-viral particles. Embryonic fibroblasts, adult retinal pigment epithelial cells and adult testicular cells were tested for their ability to de-differentiate and acquire ESC-like properties. The level of reprogramming was determined by cellular morphology, presence of surface markers and ultimately, the expression of pluripotency marker genes.

### Results

#### *Morphology and cell surface marker analysis of transduced human somatic cells*

Human embryonic fibroblasts (HEF) were infected with recombinant lentivirus containing the cDNAs for five transcription factors (Oct4, Sox2, cMyc, Klf4 and Nanog) to verify that previously reported results<sup>1,2</sup> could be replicated in this study. Additionally, human adult retinal pigment epithelial (RPE) cells were also infected with the same combination of lentiviruses to determine if adult human somatic cells of a different tissue origin could be effectively reprogrammed. Colonies resembling stem cell colonies were detected after 14 (RPE) or 19 (HEF) days (Figure 1). Twenty of the RPE and three of the HEF colonies were successfully expanded. While the RPE colonies could be treated with trypsin at an early stage in expansion, the HEF colonies had to be picked manually. During expansion, 10% of the RPE colonies and 80% of the HEF colonies developed fibroblastic morphology. Taken together, these suggest a more sustained and robust change in the RPE colonies as compared to HEF. Cell surface marker analysis of the colonies revealed that the RPE cells did not stain for any embryonic stem cell surface marker, whereas the HEF colonies stained positive for the hESC surface markers SSEA4, Tra1-81 and Tra1-60 (Figure 1).

p1

[Download this technical note](#)



### Gene expression analysis of transformed cells

Multiplex gene expression profiling was used to measure the relative expression of each pluripotency marker in experimental cells. HEFs did not show any expression of pluripotent markers before infection (Figure 2). Expression of the five transduced transcription factors was first detected at seven days postinfection, accompanied by a slight upregulation of the fibroblast markers Thy1 and Col5A2, the mesoderm marker Brachyury and the germcell-specific marker cRET. Earlier reports of reprogramming in adult mouse tail fibroblasts suggested activation of ALPL and a downregulation of Thy1 as well as other fibroblast differentiation marker genes.<sup>6,7,8</sup> In this experiment however, the steepest increase of ALPL was detected after day 17. Moreover, the fibroblast marker genes Thy1 and Col5A2 increased markedly at seven days postinfection, but then decreased to the control level of expression seen in untreated HEFs. These results suggest that the timeline and marker expression profiles in human and mouse fibroblasts may be different during reprogramming.

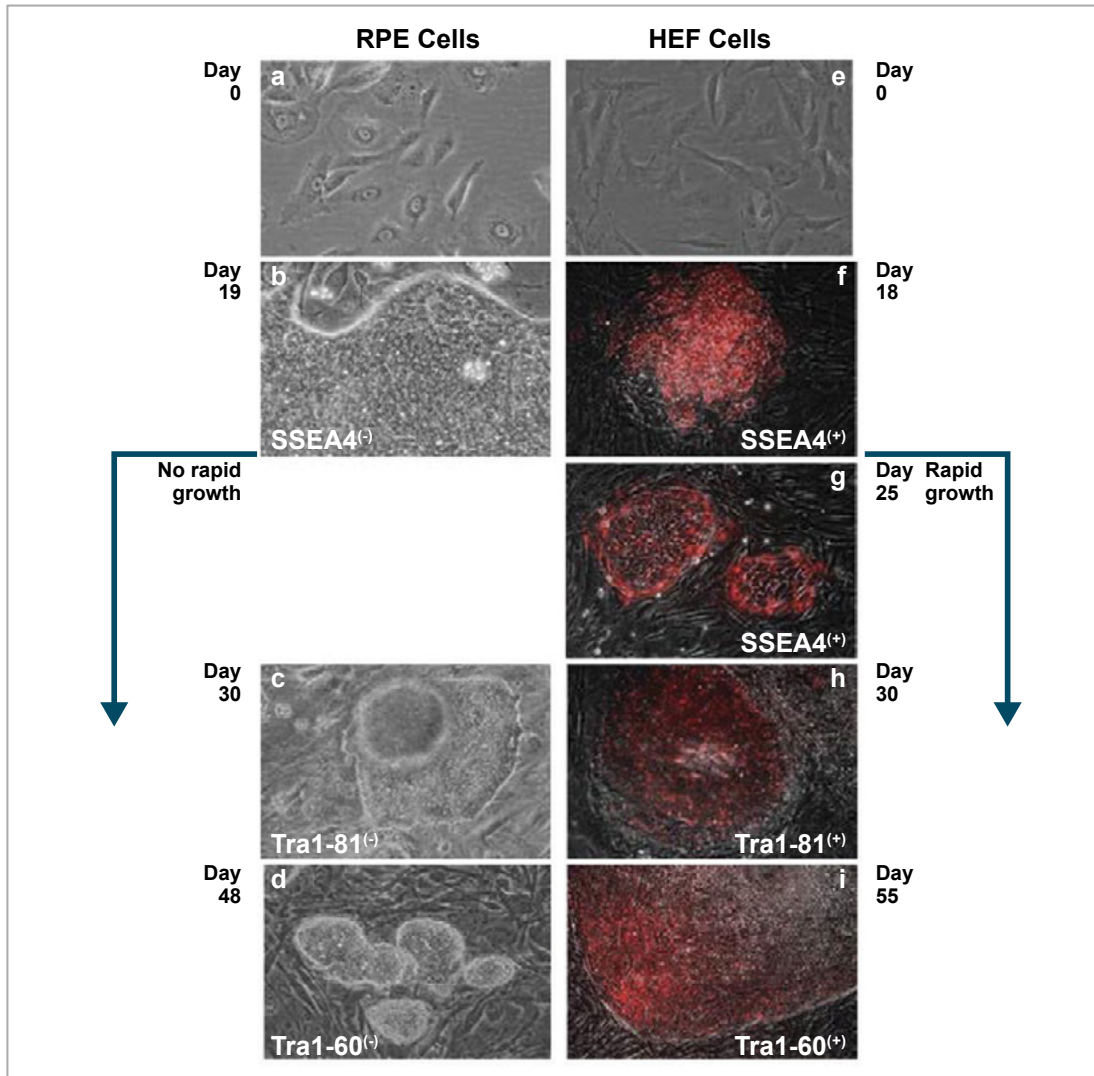
The levels of cDNA expression of pluripotent genes in the SSEA4+ population of HEFs at day 17 were greater than in the overall cell population at day seven (data not shown), suggesting a progressive enrichment of the virus-expressing cells in the SSEA4+ population. There was small but distinguishable upregulation of a number of pluripotent marker genes (hTERT, ALPL, Cripto, Sall4 and Dppa5), suggesting that within the virus-expressing SSEA4+ cells, only a small fraction of cells were de-differentiated, as described by Mikkelsen et al.<sup>8</sup> An alternative explanation would be the beginning of a global expression change leading to iPS cells. Nevertheless, after replating the SSEA4+ population, no colonies could be detected. In summary, HEF colonies expressed a complete set of endogenous pluripotent marker genes 30 days post-infection, strongly implying that these cells were de-differentiated. Further studies are necessary to determine if these cells have full pluripotentiality.

Control RPE cells infected with GFP-lentivirus had low endogenous cMyc and intermediate levels of Klf4 expression, but no expression of any other pluripotent markers (Figure 3). After transduction and clonal expansion, one RPE clone (no. 6),

had expression of Klf4, cMyc and Nanog, but lacked Oct4 and Sox2 cDNA expression. This strongly suggests that just two transcription factors, cMyc and Nanog, were transduced into these cells. For RPE colonies, endogenous Klf4 expression rose after 5-factor infection as did total (endogenous plus exogenous) Klf4 cDNA expression. However, this phenomenon was also observed in control RPE cells plated onto MEF feeder cells and grown in hESC medium (data not shown). It can therefore not be assumed that these RPE cells have been infected with the Klf4 virus, unless the cells have been tested for the viral insert. In summary, no upregulation of any of the tested pluripotent marker genes was observed 30 days post-infection in any of the resultant RPE cell lines.

In the interest of deriving iPS cells from these intermediate RPE cells, we tested whether the missing pluripotent transcription factors could be transduced at a later time point. For these experiments, RPE colonies were transduced with lentiviruses having a bi-cistronic construct containing Klf4, Oct4 or Sox2 in combination with GFP (Figure 4). Using this approach, it was possible to infect the RPE colonies, sort for GFP-positive cells to ensure that only infected cells were plated, and then infect again, thereby increasing the probability that all three transcription factors entered the cell. Gene expression analysis (Figure 3) showed that re-infection led to the uptake of Oct4, Sox2 and Klf4 cDNA. However, there was no change in SSEA4-staining (data not shown) or endogenous pluripotent marker expression (Figure 3), suggesting that sequential infection under these conditions did not lead to an iPS state in the observed timeframe of 30 days. Similarly, it has been reported that ectopic expression of Oct4 and cMyc is sufficient to induce hESC-like morphological changes in human fibroblasts without inducing pluripotency.<sup>9</sup>

Interestingly, cRET and Brachyury are up-regulated early in the reprogramming process in HEF iPS-like cells (Figure 2), while there is little to no activation of those genes after initial infection of RPE cells (Figure 3). Only after reinfection of the Nanog- and cMyc-expressing RPE cells with the missing transcription factors is there a minor increase in the expression of cRET and activation of Brachyury. One could speculate that the delayed activation in expression of Brachyury and cRET may be the cause for the failed de-differentiation in RPE cells.

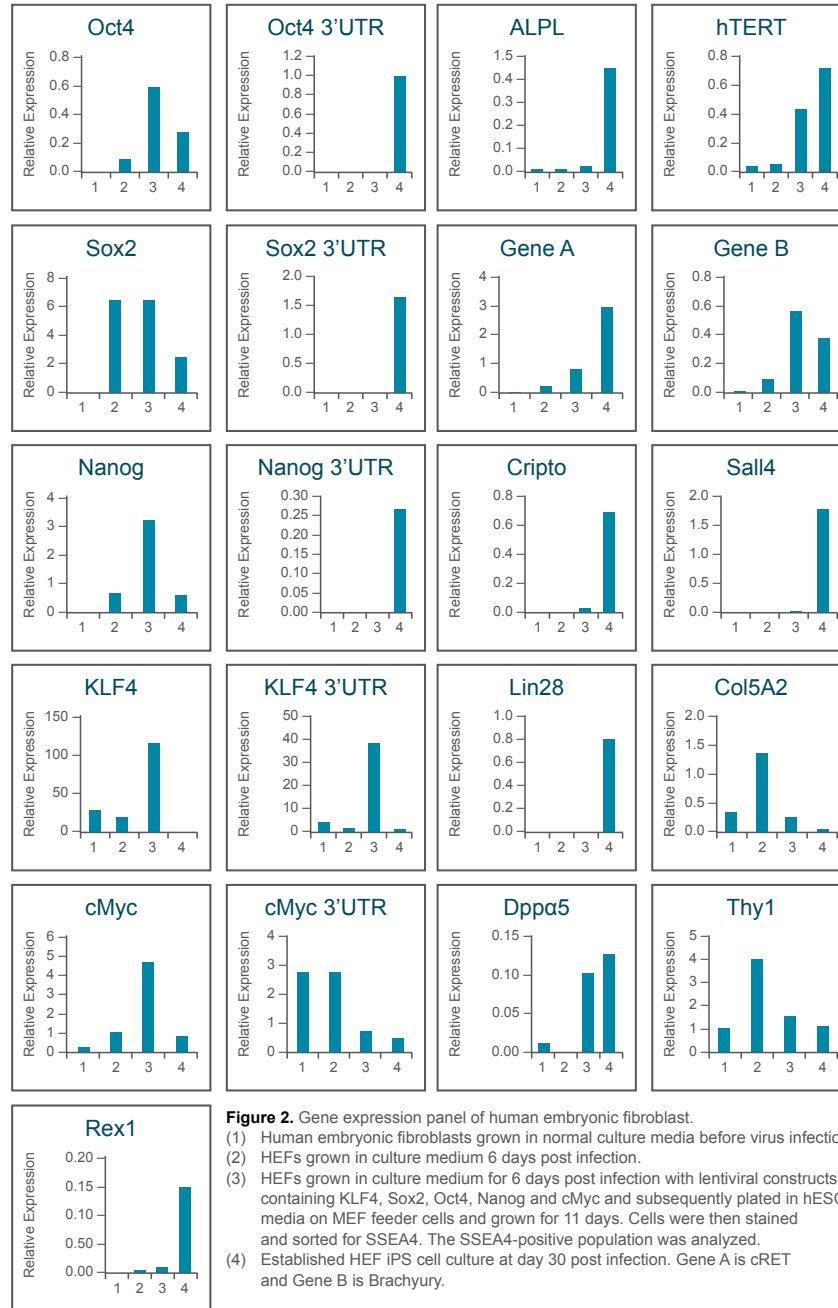


**Figure 1.** Derivation of intermediate retinal pigment epithelial cells and human embryonic fibroblast iPS-like cells.

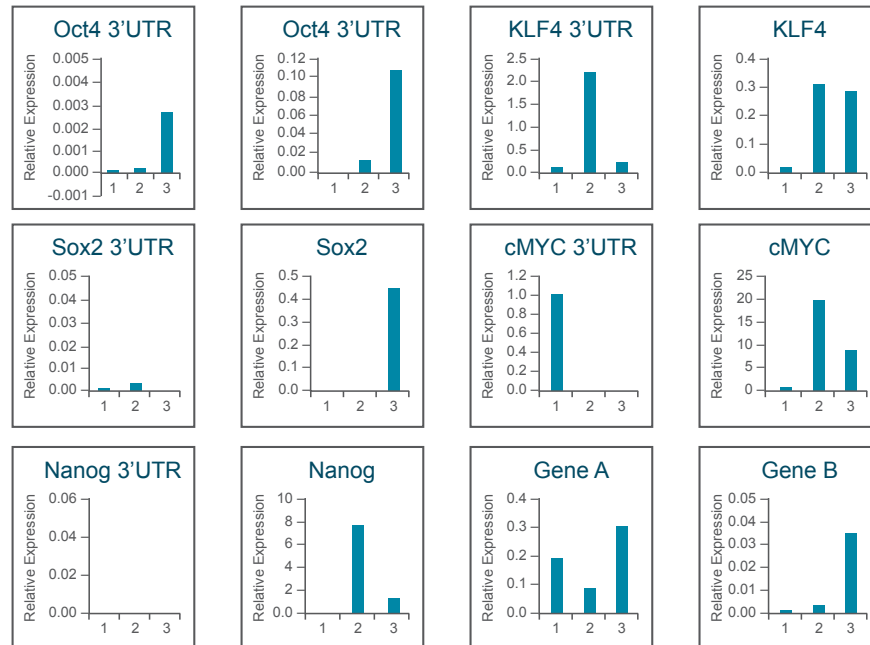
- RPE cells grown in normal media before virus infection.
- RPE cells 18 days post infection with lentivirus containing Oct4, Sox2, KLF4, cMyc and Nanog virus. Cells have been grown in RPE culture medium for 6 days on a normal culture dish and subsequently seeded onto MMC-treated MEF feeder cells at a density of  $5 \times 10^4$  cells. Colonies emerged after 14 - 18 days with a frequency of approximately 1/500. Colonies did not stain for SSEA4.
- The colonies could be picked and passaged onto new feeder cells by manual picking. The resulting colonies maintained their morphology, but grew slowly. Colonies did not stain for Tra1-81.
- RPE colonies were trypsinized and passaged onto new feeder cells without losing their morphology. Colonies did not stain for Tra1-60.
- Human embryonic fibroblasts grown before virus infection. (f) Six days post infection, infected HEF cells were seeded onto MEF feeder cells at a density of  $5 \times 10^4$  cells. HEF colonies emerging 18 days post infection at a rate of 1 in  $1 \times 10^4$ . Colonies staining for SSEA4 were manually picked and passaged onto fresh feeder cells.
- Colonies grew rapidly and maintained their expression of SSEA4. (h) Resulting colonies also stained for Tra1-81 and (i) Tra1-60. The colonies maintained their morphology at a rate of 20%. Differentiated cells were observed at the shape and size of fibroblasts.

p3

[Download this technical note](#)



[Download this technical note](#)



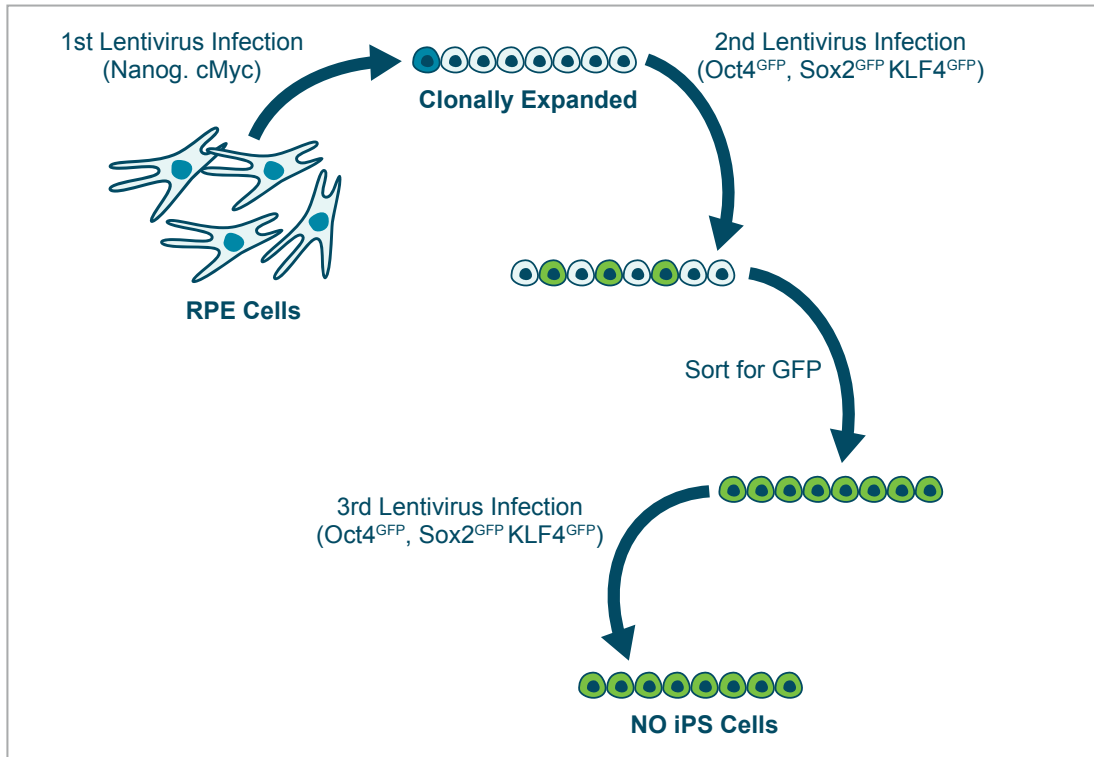
**Figure 3.** Gene expression panel of retinal pigment epithelial cells.

- (1) RPE cells grown in normal media before virus infection.
- (2) RPE cells grown on mitomycin C (MMC) treated mouse embryonic fibroblast (MEF) feeder cells in hESC media at 30 days post-infection with lentiviruses containing cDNAs encoding Oct4, Sox2, Klf4, cMyc and Nanog transcription factors.
- (3) RPE cells grown on MMC treated MEF feeder cells after two more rounds of subsequent virus.

### Non-viral delivery of transcription factor genes

Non-viral genetic delivery mechanisms are highly desirable to avoid the drawbacks of viral vectors in the clinical setting. Single walled carbon nanotubes (SWNT) have been reported to bind and deliver DNA into the nuclei of cells.<sup>10</sup> The five transcription factor plasmid DNAs (Oct4, Sox2, Klf4, cMyc and Nanog) were bound to SWNT. Human testicular fibroblasts were stably transformed with a lentivirus encoding a red fluorescent protein (RFP) driven by the Nanog promoter to create RFP reporter cells (HT42 NP). Detection of RFP production served as a visual indicator of Nanog promoter activation and the initiation of cellular reprogramming. HT42 NP cells were incubated with the plasmid-bound SWNT particles. The transfection efficiency achieved with this method was consistently below 5% as determined by flow cytometry analysis (data not shown). However, five days after transfection, a few RFP-positive colonies could be observed (Figure 5). These colonies grew slowly until day 12 and then

stopped, still expressing RFP. While no cell lines resulted, it was possible to stain two RFP-positive colonies with anti-SSEA4 antibody (Figure 5). The results suggested that after transfection with 5 transcription factor plasmid DNAs, HT42 NP cells may enter a transient state of activation of both SSEA4 and the Nanog promoter. However, the failure to establish cell lines suggested that the levels of transcription factor proteins, or the timing of expression, were insufficient to allow stable reprogramming. Published results with a doxycycline-inducible retrovirus<sup>6,7,8</sup> suggest the necessity of stable transcription factor expression over at least nine days in mouse cells. Therefore, the expression profile of HT42 NP cells was evaluated after delivery of the five transcription factor plasmid DNAs coupled to SWNT. RNA samples were collected at Day 0 (8 hours), 7 and 14 days after treatment and expression was analyzed using multiplex XP-PCR. Expression of plasmid DNAs introduced by SWNT was confirmed by XP-PCR and, as expected, the expression levels



**Figure 4.** Flow chart for re-infection experiment.

drastically decreased from day zero to day seven (Figure 6). Interestingly, expression levels of two pluripotent marker genes in the panel, *Sall4* and endogenous *Nanog* (*Nanog* 3'UTR), increased coincident with the expression from the plasmid DNA. Thus, a transient activation of the *Nanog* promoter in HT42 NP cells was confirmed by two different methods. Surprisingly, expression levels of the reported early reprogramming marker ALPL<sup>6,7,8</sup> were decreased relative to controls. This could be due to the fact that HT42 NP cells already have a high level of ALPL expression and in this case, the expected increase in ALPL may not have been detectable against the high background. In further support of the induced changes in human testicular (HT) cells after delivery of five transcription factors DNAs, *cRET* and *Brachyury* were consistently upregulated. Combined with similar findings in virally-transduced HEF cells, there is

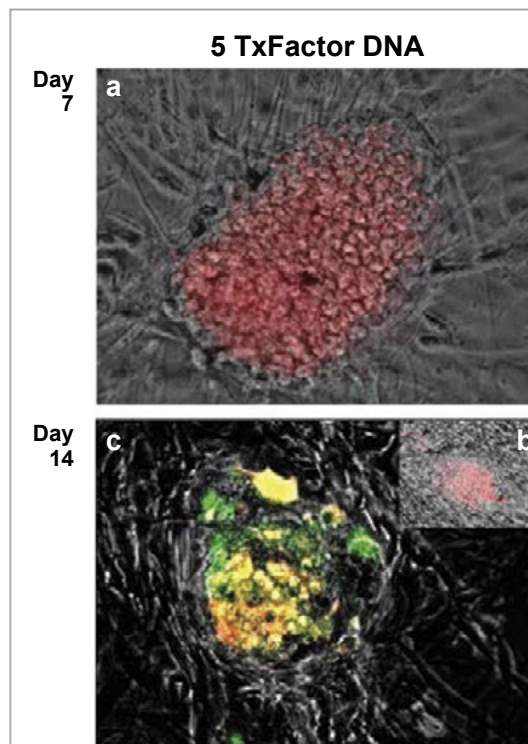
additional support for the idea that HT cells are responsive to reprogramming signals, but either the timing and/or co-expressed levels of transduced genes are not effective to induce a stable expression of pluripotent genes. Alternatively, *Sox2* expression could not be detected in the treated HT cells and this could account for the observed defective reprogramming.

To investigate whether the failure to reprogram might be related to the type of delivery particle, HEFs were treated with the stable aggregates formed between the five transcription factor plasmid DNAs and polyethylene imine (PEI) particles. The latter delivery method consistently produced transfection efficiencies >25% (data not shown), in comparison to less than 5% with SWNT. The expression levels of the five transcription factors increased for 24 hours after infection and then dropped after 72 hours (Figure 7). In contrast to the SWNT treatment of HT42 NP cells, no change

was observed in expression of Sall4 or endogenous Nanog in HEF. Additionally, no colonies were observed with HEF (data not shown). The differences in pluripotent marker expression between HT and HEF imply that perhaps a more rapid activation of Nanog may occur in the HT cells than in HEFs, which may account for the presence of HT colonies and absence of HEF colonies with particle-based transduction. It is possible that intrinsic epigenetic genomic plasticity may play a role determining the potentiality of different cell types for iPS cell reprogramming. Additionally, the same two genes (cRET and Brachyury) that were upregulated in the HT experiments above were also transiently activated in HEFs (Figures 6 and 7, respectively), suggesting that these genes may be early downstream targets of one or a combination of the transduced cDNAs during the reprogramming process.

### Discussion

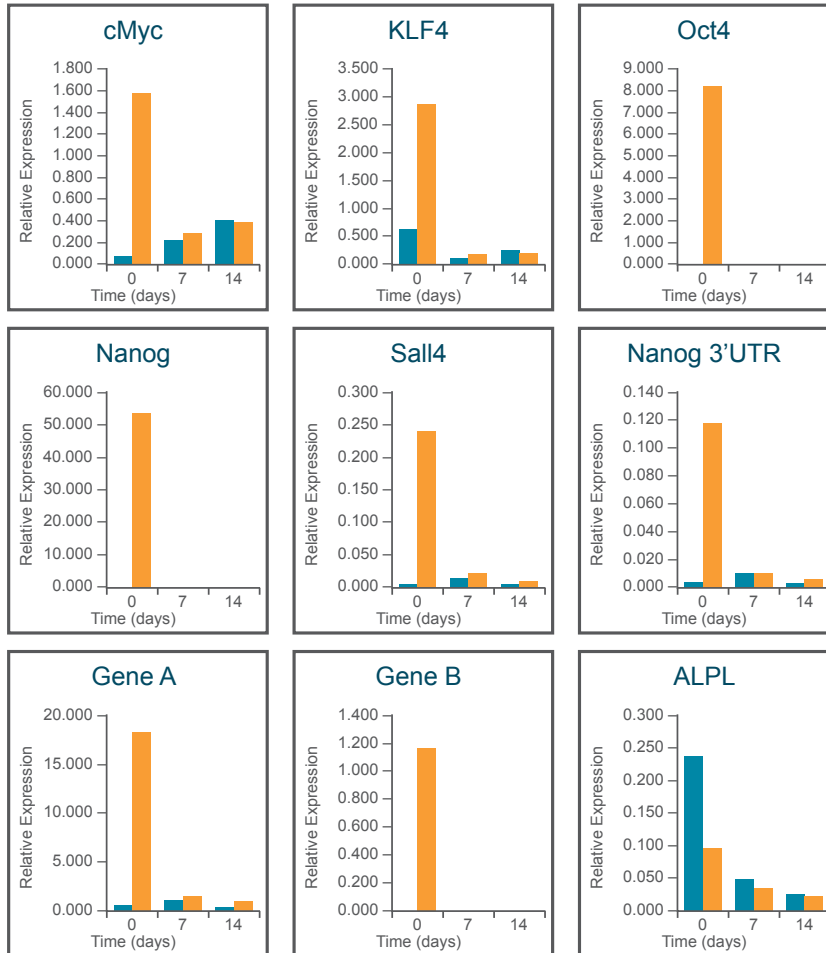
The ability to generate iPS cells from patient somatic cells promises regenerative therapy without the need for immunosuppressive drugs to prevent unmatched cell rejection. However technical and safety obstacles must be overcome before this promise becomes reality. Previous work demonstrated the ability to create iPS cells from both embryonic and adult fibroblasts employing retroviral vectors.<sup>1,2</sup> In this study, we demonstrate the effective delivery of pluripotency genes in both adult human fibroblasts as well as retinal epithelial cells using retroviral vectors. ESC-like colonies were derived from human embryonic fibroblasts (HEF) that were infected with recombinant lentiviruses containing the coding sequences for five transcription factors (Oct4, Sox2, cMyc, Klf4 and Nanog). Expression of a complete set of endogenous, pluripotent marker genes was detected at day 30, strongly implying that these cells were de-differentiated. Further studies are underway to determine if they were truly reprogrammed to the ESC state. A non-pluripotent stem cell line, derived from retinal pigment epithelial (RPE) cells, grew in hESC-like colonies by expressing only Nanog, cMyc and Klf4. Re-infection of this intermediate RPE cell line with the missing transcription factors (Oct4 and Sox2), however, was not sufficient to induce pluripotency.



**Figure 5.** SWNT-bound 5 transcription factor DNA leads to colony formation without establishing a cell line.

- (1) HT42 NP cells 7 days after transfection with SWNT bound 5 Factor DNA. RFP expression is driven by the Nanog promoter.
- (b) HT42 NP cells 14 days after transfection with SWNT bound 5 Factor DNA. RFP expression is driven by the Nanog promoter.
- (c) HT42 NP cells 14 days after transfection with SWNT bound 5 Factor DNA. RFP expression is driven by the Nanog promoter. Colonies were stained with a FITC coupled SSEA4 antibody. A merge of the pictures lead to yellow color in Nanog and SSEA4+ cells. There was no further growth detectable after day 12 of transfection. All pictures are 20 x magnifications.



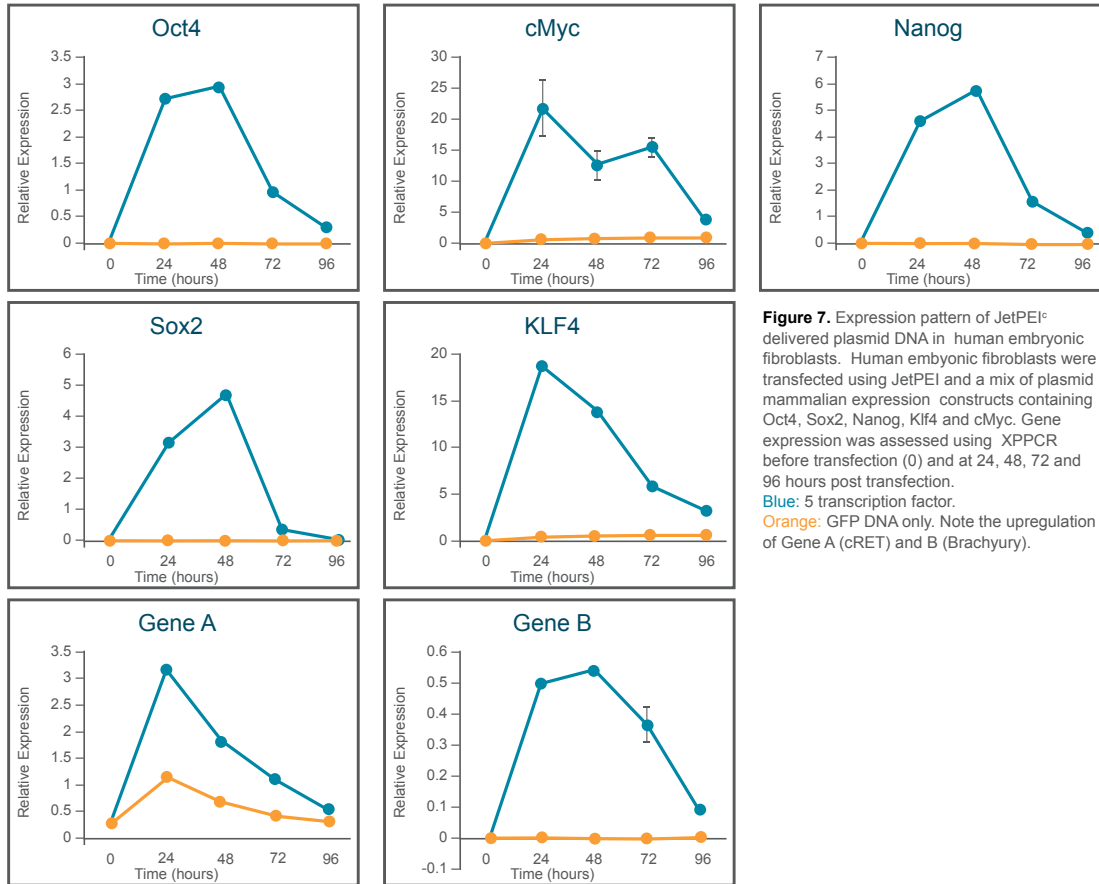


**Figure 6.** Gene expression analysis of HT42 NP cells treated with SWNT-bound 5-Factor DNA. Either 5-factor expression plasmid DNA coding for Oct4, Sox2, KLF4, cMyc and Nanog were bound onto SWNT (orange bars) or SWNT were left untreated (blue bars). RNA samples were taken from the total cell population 8h (0), 7 days (7) and 14 days (14) post transfection and assayed via XP-PCR. Sox2 expression could not be detected. Gene A is cRET and Gene B is Brachyury.

Recent work demonstrated that only Oct4 and Sox2 were essential for the reprogramming of embryonic and adult fibroblasts<sup>11</sup> and other factors, such as cMyc, Klf4, Nanog and Lin28, serve to complement and enhance the efficiency of colony formation.<sup>4,11</sup>

A major obstacle to the use of iPS cells in regenerative therapy is the retroviral vectors used to induce pluripotency. These genome-integrating viruses have the potential to activate proto-oncogenes and generate tumors. In this study, we demonstrate the first steps toward a technology of a non-viral, particle-based

gene delivery mechanism with the aim of reprogramming human somatic cells. Single-walled carbon nanotubes (SWNT) were used to transfer a mixture of the five transcription factor cDNA plasmids into primary HT cells stably transfected with a Nanog-RFP reporter construct (HT42 NP). The transfected HT42 NP cells demonstrated transient colony formation, endogenous Nanog activation and expression of two suspected indicators of reprogramming, cRET and Brachyury. It is hypothesized that HT cells may have either an epigenetic genomic plasticity, or intrinsic controls, that account for their apparent ease of transient



**Figure 7.** Expression pattern of JetPEI<sup>L</sup>-delivered plasmid DNA in human embryonic fibroblasts. Human embryonic fibroblasts were transfected using JetPEI and a mix of plasmid mammalian expression constructs containing Oct4, Sox2, Nanog, Klf4 and cMyc. Gene expression was assessed using XPPCR before transfection (0) and at 24, 48, 72 and 96 hours post transfection. **Blue:** 5 transcription factor. **Orange:** GFP DNA only. Note the upregulation of Gene A (cRET) and B (Brachyury).

colony formation, Nanog activation and marker gene expression. The transient expression profile of HEFs was determined after introduction of the five transcription factor plasmid cDNAs using PEI particles. Here the same suspected marker genes of reprogramming were transiently upregulated, while no colony formation could be observed within 30 days. Transduction can lead to a heterogeneous population of cells with various plasmid copy number and cells that receive some, but not all the transcription factors. This may account for the uncoordinated timing and expression levels for the five co-transduced genes, resulting in only a few reprogrammable cells.

While the overall efficiency of gene transduction was low, these studies indicate that somatic cells can be effectively reprogrammed to express pluripotency markers with retroviral gene delivery methods. It is also possible to observe minute changes in gene expression after a transient transfection protocol using SWNT. Although there is still much to be done in order to effectively reprogram adult somatic cells for application in the clinical environment, this work demonstrates that the expectation of generating iPS cells for therapeutic purposes is well on its way.

### Materials and Methods

**Lentivirus transduction:** Lentivirus production was performed as described.<sup>12</sup> Retinal pigment epithelial cells (RPE) and human embryonic fibroblasts (HEF) were infected with lentiviral particles containing the cDNA of Oct4, Sox2, Klf4, cMyc and Nanog at an approximate MOI of 10. Infected cells were grown under normal conditions in untreated dishes for six days and subsequently seeded onto mouse embryonic fibroblasts (MEF) feeder cells in hESC medium at a density of  $5 \times 10^4$  per 10-cm dish. Colonies were picked and clonally expanded with passaging every 3-7 days, onto fresh MEF feeder cells by either trypsinization (RPE) or manual picking and eventual passing using collagenase IV (HEF).

**Cell surface marker staining:** Rabbit-antihuman Tra1-60, Tra1-81 (Chemicon<sup>®</sup>) and SSEA4 (eBioscience<sup>®</sup>) antibodies were incubated with live cells for one hour in normal growth medium followed by washing and detection of the primary antibody with a secondary, TRITC-labeled sheep-anti-rabbit antibody.

**Gene expression profiling:** Forward and reverse primers for a 25 gene multiplex were constructed. Separate primers were designed to the coding sequence (CDS) and the 3 untranslated regions (3UTR) of Oct4, Sox2, Nanog, Klf4, cMyc of the same mRNAs to distinguish between endogenous plus exogenous expression (CDS) and endogenous expression only (3UTR). Additional primer sets were designed for Lin28, Col5A2, mouse GAPDH (to test for feeder layer contamination), human GAPDH, Gene A, Gene B, TERT, Thy1, Rex1 (aka ZFP42), Dppa5, ALPL, beta-Actin, Sall4 and Cripto (aka TDGF1). Primers for the internal control Kanamycin resistance gene (KanR) was also included in the multiplex design. The multiplex reverse transcription-PCR reaction for express profiling (XP-PCR) was performed according to the manufacturer's instructions using the GenomeLab™ GeXP Start Kit (SCIEX), separated by capillary electrophoresis and analyzed in the GenomeLab™ GeXP Genetic Analysis System (SCIEX, Figure 8). Gene expression data was normalized to the internal control gene (KanR), quantified using a standard curve of hESC RNA ranging from 1ng to 500ng, and evaluated with GenomeLab™ GeXP Quant Tool software (available for download on SCIEX website) using hGAPDH and beta-Actin as normalization genes.

**DNA fingerprinting:** Fingerprinting was done by Cell Line Genetics Inc.\* Both of the derived cell lines (RPE and HEF iPS-like) are being grown continuously and have been confirmed to be identical to the original cell line by DNA fingerprinting.

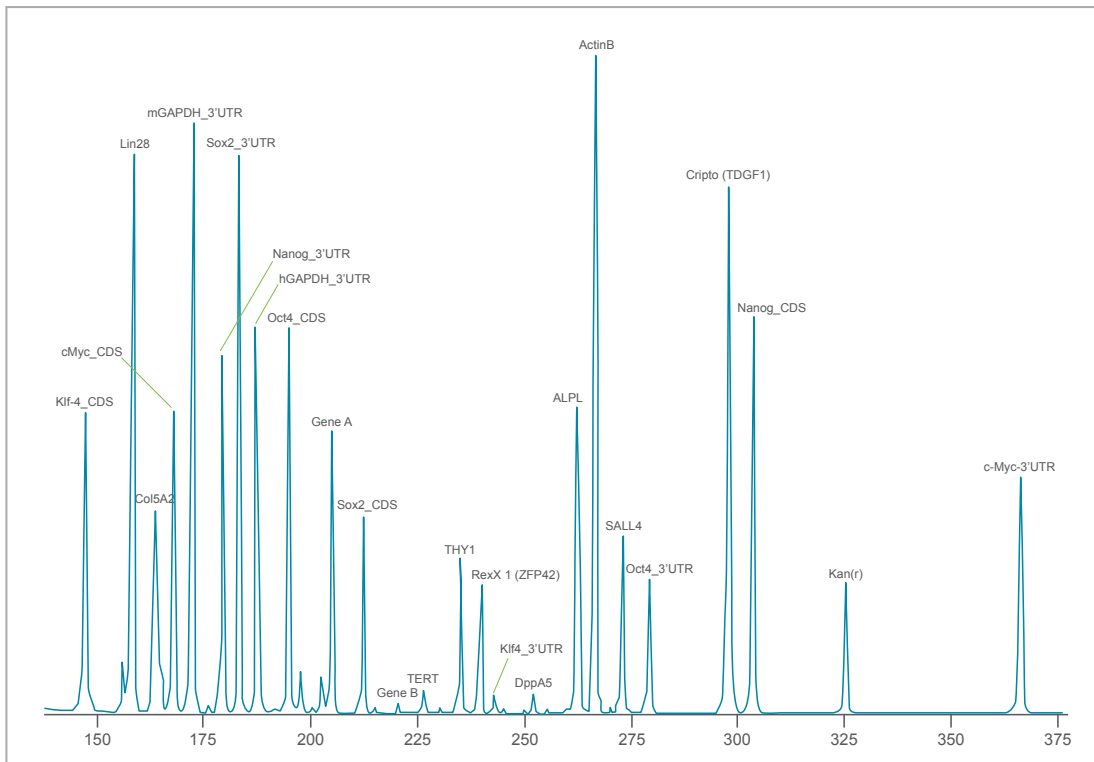
**HT42 NP cells:** Human testicular (HT) cells were prepared as described (PCT/US2005/ 047437; WO2006/074075). In short, HT42 fibroblastic cells from adult human testicular tissue were recovered after enzymatic digestion and selection from adherent cells. RFP reporter cells (HT42 NP) were constructed by transducing HT42 cells with a lentivirus containing a Nanog promoter driving the expression red fluorescent protein (RFP).

**SWNT binding and treatment:** DNA binding was performed as described.<sup>10,13,14</sup> SWNT were diluted to a final concentration of 1ug/mL and added to  $1 \times 10^6$  HT target cells followed by incubation for two hours in the presence of 100 mM Chloroquine. Cells were washed and incubated with normal growth medium overnight. Cells were then trypsinized and plated onto MEF feeder cells at a density of  $1 \times 10^5$  cells per 10 cm-dish and cultivated in hESC medium.

**jetPEI™\* transfection:** PEI transfection was performed according to the manufacturer's instructions (Polyplus-transfection, Inc.).

Gene Name	Fragment Size (nucleotides)
Kif-4_CDS	147
Lin28	158
Col5A2	164
c-Myc_CDS	168
mGAPDH_3'UTR	172
Nanog_3'UTR	179
Sox-2_3'UTR	184
hGAPDH_3'UTR	188
Oct-4_CDS	195
Gene A	205
Sox-2_CDS	213
Gene B	220
TERT	226

Gene Name	Fragment Size (nucleotides)
THY1	235
RexX 1 (ZFP42)	240
Kif-4_3'UTR	243
DPPA5	252
ALPL	262
ActinB	266
SALL4	273
Oct-4_3'UTR	279
Cripto (TDGF1)	298
Nanog_CDS	304
Kan(r)	325
c-Myc-3'UTR	366



**Figure 8.** Multiplex gene list and electropherogram. Primer sets for twenty-five gene targets (see chart at top) were assembled for XP-PCR reactions, and were then separated by capillary electrophoresis (see electropherogram). Each peak represents the PCR product of a gene. The area under each gene peak is used to quantitate the expression of that gene in each sample.

[Download this technical note](#)

### References

1. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006. 126(4):663-76.
2. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007. 131(5):861-72.
3. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*. 2007. 448(7151):318-24.
4. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA. Induced pluripotent stem cell lines derived from human somatic cells. *Science*. 2007. 318(5858):1917-20.
5. Hanna J, Wernig M, Markoulaki S, Sun CW, Meissner A, Cassady JP, Beard C, Brambrink T, Wu LC, Townes TM, Jaenisch R. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science*. 2007. 318(5858):1920-3.
6. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell*. 2008. 2(3):230-40.
7. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, Jaenisch R. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*. 2008. 2(2):151-9.
8. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A. Dissecting direct reprogramming through integrative genomic analysis. *Nature*. 2008. 454 (7200): 49-55. Erratum in: *Nature*. 2008. 454 (7205):794.
9. Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K. Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci U S A*. 2008. 105(8):2883-8.
10. Kam NW, Liu Z, Dai H. Functionalization of carbon nanotubes via cleavable disulfide bonds for efficient intracellular delivery of siRNA and potent gene silencing. *J Am Chem Soc*. 2005. 127(36):12492-3.
11. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature*. 2008. 451(7175):141-6.
12. Ramezani A, Hawley RG. Generation of HIV-1-based lentiviral vector particles. *Curr Protoc Mol Biol*. 2002. Chapter 16: Unit 16.22.
13. Singh R, Pantarotto D, McCarthy D, Chaloin O, Hoebeke J, Partidos CD, Briand JP, Prato M, Bianco A, Kostarelos K. Binding and condensation of plasmid DNA onto functionalized carbon nanotubes: toward the construction of nanotube-based gene delivery vectors. *J Am Chem Soc*. 2005. 127(12): 4388- 96.
14. Kam NW, Liu Z, Dai H. Carbon nanotubes as intracellular transporters for proteins and DNA: an investigation of the uptake mechanism and pathway. *Angew Chem Int Ed Engl*. 2006. 45(4):577-81.







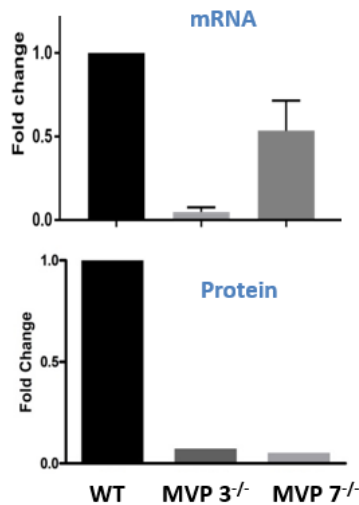
## Confirming Gene Mutation by CRISPR-Cas9 at the Protein Level and Identifying Proteome-Wide Changes

### Using SWATH® Acquisition on TripleTOF® Systems

Qifeng Lin<sup>1</sup>, Larry W.L. Low<sup>1</sup>, Stephen Tate<sup>2</sup>, Jayantha Gunaratne<sup>1,3</sup> and Thomas J. Carney<sup>1,4</sup>  
<sup>1</sup>Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A\*STAR), Singapore; <sup>2</sup>SCIEX, Canada; <sup>3</sup>Yong Loo Lin School of Medicine, National University of Singapore, Singapore, <sup>4</sup>Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Genome editing tools are invaluable to life science researchers, allowing changes to be made to an organism's DNA and then studying the effects to gain biological insights. The most recent gene editing tool is CRISPR-Cas9 which has rapidly gained adoption as it is faster, simpler and cheaper than previous methods. To ensure fidelity of biological conclusions, it is important to confirm the success of any gene editing at the protein level. Immunoblotting is typically used; however, this strategy can be limited by the availability and quality of antibodies.

Here, a label-free mass spectrometry approach has been applied both to confirm protein level changes induced by gene editing and to measure the proteome wide changes that occur in



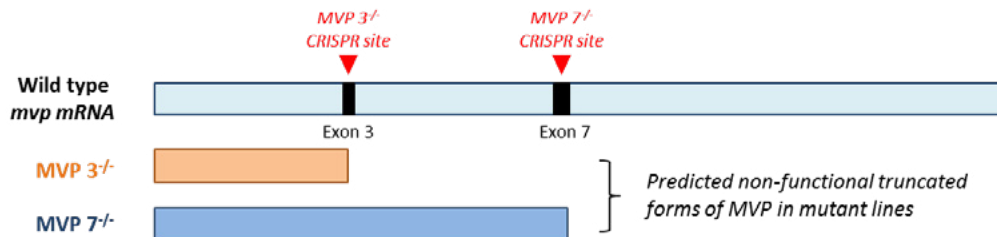
**Figure 1. Ensuring Gene Editing and Loss at the Protein Level.** qPCR analysis indicates total loss of mutant *mvp* mRNA in the first CRISPR allele (MVP3<sup>-/-</sup>) but residual mutant mRNA in the second allele (MVP7<sup>-/-</sup>). No antibodies are available for this protein in zebrafish therefore SWATH acquisition was used to confirm that gene editing was indeed effective and a significant reduction in the amount of MVP was observed at the protein level for both knockout alleles.

response to the system perturbation in a single experiment. Additionally, the method does not require antibodies, making it broadly applicable to many organisms.

To demonstrate feasibility, CRISPR-Cas9 was used to knock out the Major Vault Protein (MVP) in zebrafish, and SWATH Acquisition was used to confirm the protein knockdown (Figure 1) and to study the proteome-wide changes that were induced (Figure 4).

### Key Feature of SWATH Acquisition in Combination with CRISPR-CAS9 Studies

- SWATH® Acquisition<sup>1</sup>, a data independent acquisition strategy, is the method of choice due to the comprehensive nature of the data that allows both high quality protein level quantitation as well as data re-interrogation for specific peptides to proteins of interest, without the need to rerun the sample.
- No antibodies are required, making the method more broadly and readily applicable to any protein and any organism.
- Provides a higher throughput workflow for screening gene editing experiments, allowing direct protein measurement.



**Figure 2. Mutant Alleles Generated by CRISPR-Induced Frameshift.** Two mutations were made in the Major Vault Protein (MVP) in the Zebrafish model system, one in exon 3 (MVP 3<sup>-/-</sup>) and one in exon 7 (MVP 7<sup>-/-</sup>). Embryos were grown and studied however no significant change in phenotype was observed.

## Methods

**Sample Preparation:** Exons 3 and 7 of the zebrafish (*Danio rerio*) *mvp* gene were targeted using CRISPR-Cas9<sup>2</sup> (Figure 2). Three biological replicates each of de yol ked zebrafish 30 hours post fertilization embryos (300 per replicate) were prepared for wild type and *mvp* mutants. Samples were lysed using 50mM Tris-HCl, 100mM NaCl, 2% sodium deoxycholate. Proteins were reduced, alkylated, and digested first with Lys-C followed by trypsin. Resulting peptides were then cleaned-up using reverse phase chromatography.

**Chromatography:** Separation of the digested samples was performed on a NanoLC™ 425 System (SCIEX) operating in trap elute mode. Peptides were first trapped on a 350µm x 0.5mm trap, then eluted onto a 75µm x 15cm column (both packed with ChromXP™ column (SCIEX)). Elution was done with a long two-phase gradient (10-18% solvent B in 55 min, then 18-30% in next 60 mins). Total protein injected on column was 2 µg.

**Mass Spectrometry:** The MS analysis was performed on a TripleTOF® 6600 system (SCIEX) using a NanoSpray® Source III. Data dependent acquisition (DDA) was performed using 50 MS/MS per cycle with 50 msec accumulation time. Variable window SWATH® Acquisition methods were built using 120 variable windows with 25 msec accumulation time.

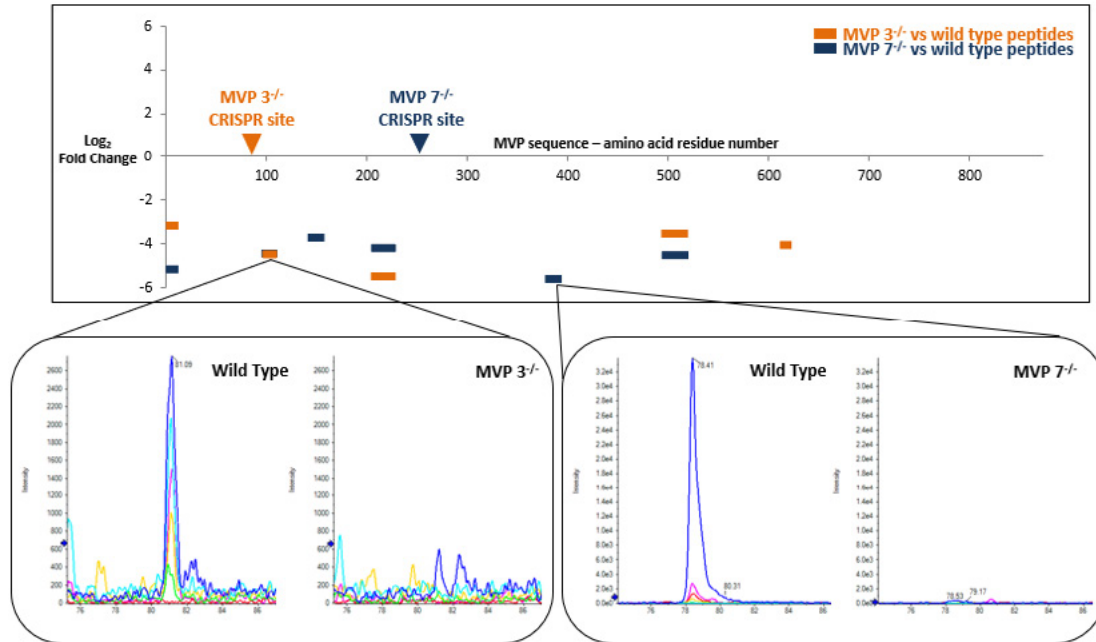
**Data Processing:** A spectral library was created by processing 3 DDA runs from a pool of wild type samples with ProteinPilot™ 5.0 Software using a thorough search. SWATH data was processed using the OneOmics™ platform in the SCIEX Cloud<sup>3</sup>. Manual data processing was performed using PeakView® Software.

## Confirming Protein Level Mutation of the Edited Gene

Two different mutant indel alleles were generated for this study by a CRISPR-induced frameshift and were predicted to generate two non-functional truncated forms of MVP in the mutant lines; MVP 3<sup>-/-</sup> and MVP 7<sup>-/-</sup> (Figure 2). Observation of the mutant embryos revealed no significant phenotype change in either mutant. Additionally, mRNA analysis indicated significant mRNA decay of one allele, but not the other (MVP 7<sup>-/-</sup>) mutant (Figure 1, top). The lack of an observable change in phenotype in addition to the ambiguous mRNA data, made protein-level validation critical.

Typically, protein-level validation would be accomplished using an antibody-based technique such as western blot or ELISA. However, in this case, no antibody was available for zebrafish MVP. Instead, protein level analysis was performed by mass spectrometry using SWATH acquisition for quantitation of all proteins within the mutant lines. The results indicated that MVP protein levels were, in fact, negligible for *both* mutant lines (Figure 1, bottom).

Further scrutiny of the SWATH data indicated loss of peptides across the entire MVP protein sequence. Within the SWATH acquisition, 7 peptides across the entire MVP sequence were compared between the wild type and two mutant alleles. All 9 peptides were of significantly lower abundance within the mutants relative to the wild type (Figure 3). Even peptides within portions of the mutant MVP non-truncated sequences were also found to be of significantly lower abundance. This confirms that the CRISPR-Cas9 editing successfully targeted the gene and resulted in the loss of the entire MVP protein, and not just the portion located beyond the CRISPR site. This indicates instability of the truncated mutant MVP protein of both alleles, and thus abolition of the whole protein had been accomplished.



**Figure 3. Measured MVP Peptide Signals Across the Full Protein Sequence.** Extraction of the 7 peptides for MVP from the SWATH acquisition data allows for quantitation of the peptide level changes between the 9 samples. Comparing the peptide areas to those observed in the wild type samples using Browser shows that all peptides across the protein sequence were of significantly lower abundance relative to wild type. This confirms that the CRISPR-Cas9 editing successfully knocked down the protein, resulting in loss of the entire MVP protein (not just post CRISPR site). This provides more direct measurement of knock down success than mRNA experiment (Figure 1). Examples of the raw XIC data for 2 peptides is shown (bottom).

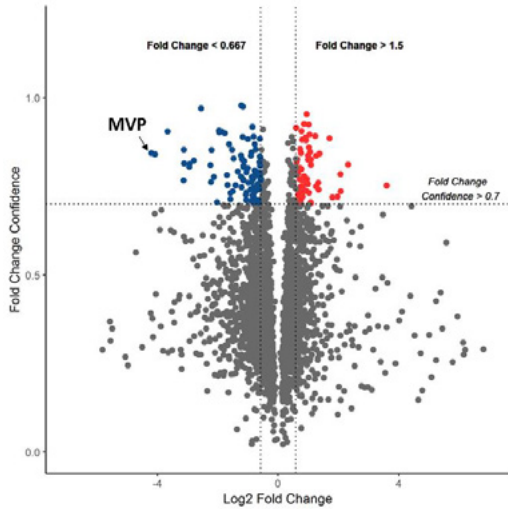
### A Global View of All Proteins Impacted by the Edited Gene

A huge advantage of SWATH Acquisition for confirmation of genome editing at the protein level is that large numbers of proteins are analyzed in each experiment and not simply the protein product of the gene editing event. This allows a comprehensive picture of any biological changes the editing may induce. In this study, about 3800 proteins were quantified across the 6 samples (3 biological replicates each for wild type and mutant). As shown in Figure 4, within the MVP 7<sup>-/-</sup> mutant, a subset of 65 proteins were significantly over expressed while 93 proteins were under expressed compared with the wild type.

A Venn diagram comparison of the number of proteins that were significantly changed within each mutant is shown in Figure 5. Only a small number of the same proteins were up-regulated (11) and down-regulated (18) for both mutants, and the number of unique up-regulated proteins and unique down-regulated proteins was higher for the MVP 7<sup>-/-</sup> mutant.

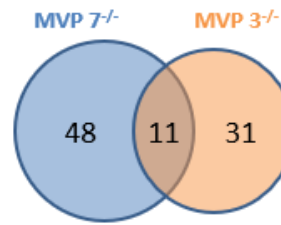
### Investigating the Induced Biology from Gene Knockdown

To gain further knowledge of the impact of the gene editing event, the proteins that were significantly up- and down-regulated were more closely investigated. Figure 6 shows the biological process information associated with the proteins over- and under-expressed for the MVP 7<sup>-/-</sup> mutant. While the MVP protein currently has no known function, proteins belonging to multiple groups of related processes are affected by the gene editing including cell adhesion, inflammation, and cytoskeleton processes. However, there is emerging evidence that MVP may play a role in infection or stress responses. This could explain why minimal changes were observed in the mutant phenotypes as there was no cellular stress invoked in the zebrafish prior to sample isolation.

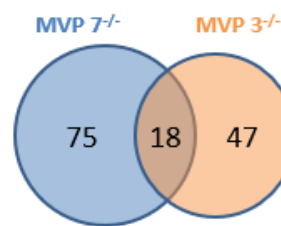


**Figure 4. Proteome Wide Changes Measured by SWATH Acquisition.** About 3800 proteins were quantified in the 6 samples. In the volcano plot shown here for log fold change vs protein confidence for MVP 7<sup>-/-</sup> vs wild type. Using a filter of >1.5-fold change, >0.7 confidence and ≥2 peptides, there were 42 proteins increased and 65 proteins decreased in expression.

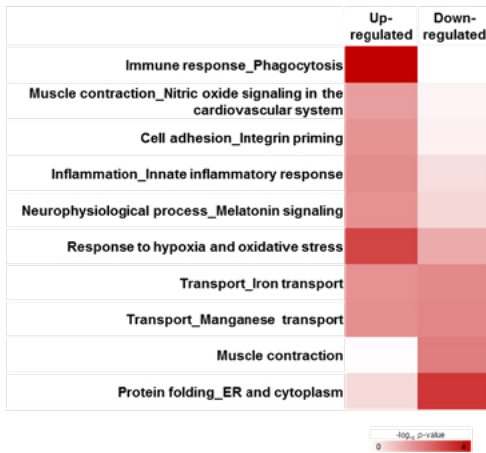
**Up-Regulated Proteins**



**Down-Regulated Proteins**



**Figure 5. Comparison of Significantly Up and Down Regulated Proteins.** Using a filter of >1.5-fold change, >0.7 confidence and ≥2 peptides, a total of 90 proteins were up-regulated for both mutants with only 11 of those in common between the two alleles. Conversely, a total of 140 proteins were down-regulated for both mutants with only 18 of those in common.



**Figure 6. Biological Process Network Functional Enrichment.** Top ten biological process networks significantly enriched among differentially regulated proteins in MVP 7<sup>-/-</sup> mutant. Differentially regulated zebrafish proteins were mapped to their respective human orthologues before functional enrichment analysis using Metacore (Clarivate Analytics). MVP proteins currently have no known function however proteins belonging to several different groups of related processes are significantly affected. Overall, lack of a coherent pathway modulation supports the “Sentinel hypothesis” for Vaults, a new emerging theory that MVP is responsible for inflammatory or stress response.

**Conclusions**

Gene editing techniques have come a long way. With the advent of the CRISPR-Cas9 technique, genes can now be edited with incredible accuracy. mRNA experiments are typically used to verify gene editing but often mRNA and protein levels do not correlate. Thus, confirmation at the protein level is absolutely critical. With SWATH Acquisition, not only is protein level verification easily accomplished, but the additional advantages and information that SWATH provides makes this technique ideally suited for deeply exploring any organism after gene editing experiments.

- SWATH Acquisition verifies gene editing at the protein level (mutations, knockouts, truncations, etc.) without the need for antibodies
  - No costly delays waiting for antibody production
  - Easily applied across diverse organisms
  - Fast experimental protocol allows research to move more rapidly
  - Single LC-MS method for investigating all samples
- Besides verifying the protein product of the gene editing event, additional information is obtained around global changes in biology that are induced
  - Global quantitative profiling measures large numbers of proteins, including additional protein expression differences due to the effects of the gene editing
  - Linked ontology information to all affected proteins provides further understanding of the biological implications of the edited gene
- SWATH Acquisition is rapid and comprehensive
  - A digital archive is created. No need to re-run samples when new questions arise
  - Can be used for high throughput profiling for screening arrays of CRISPR experiments.

### References

1. Evolution of SWATH® Acquisition Provides Large Gains in Quantified Proteins, SCIEX Technical notes RUO-MKT-02-5772-A.
2. Tracking genome-editing and associated molecular perturbations by SWATH mass spectrometry. Q. Lin, L.W.L. Low, A. Lau, H.L.F. Swa, E.W.L. Chua, Y. Matsuoka, A. Monteiro, S. Tate, J. Gunaratne, T.J. Carney. *Manuscript in preparation*.
3. OneOmics™ Project in the Cloud - Getting Started with the Proteomics Cloud ToolKit. SCIEX Technical note RUO-MKT-02-6969-B.

AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-9247-A



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





# Analysis of oligonucleotide therapeutics



## Drug Discovery and Development

# High Resolution Analysis of Synthetic Oligonucleotides

## Reverse Phase Ion Pairing LC-MS Analysis Using the X500B QTOF System

Esme Candish, Sean M. McCarthy

SCIEX, Framingham MA, USA

Oligonucleotide therapeutics and gene therapies are rapidly gaining attention as novel therapeutics. The potential disease areas for which oligonucleotide based therapies are being developed is rapidly expanding. In particular, these therapies hold high promise due to their specificity and the high potential for treating previously undruggable targets.

A defining characteristic of oligonucleotides, when compared to more commonly used protein based biotherapeutics, is that they are produced synthetically. Synthetic processes have been highly refined to enable production of high quality synthetic products with limited failure sequences or other associated impurities. Even when considering the high efficiency of modern synthetic processes there is still a need to verify oligonucleotide mass, sequence, and purity.

For the purpose of high throughput oligonucleotide analysis to verify purity and mass confirmation a common data processing solution is ProMass from Novatia LLC. This software enables rapid verification of molecular mass and assessment of purity. In addition, ProMass can annotate failure sequences and synthetic impurities based on the input of the target sequence and selection of search parameters.



Demonstrated in this technical note is the use of the SCIEX X500B QTOF System for the analysis of oligonucleotides using ion-pairing reverse phase chromatography. Highly efficient chromatographic separation with high performance negative ion mode performance produce results for unambiguous confirmation of sequences with minimal adduct formation or analysis induced modifications. Also discussed is the processing of this data using ProMass HR for automated verification of molecular mass and purity assessment.

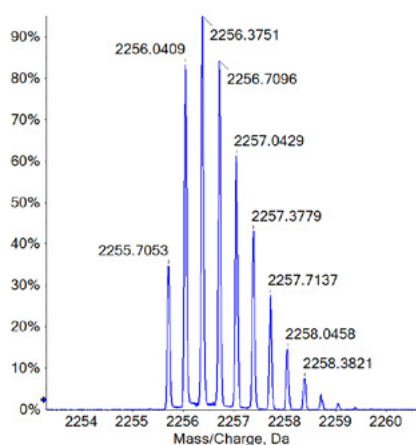


Figure 1. Isotopically resolved charge state for oligonucleotide CATGGTCTGCTGGAGGTCGTG

### Key Feature of X500B QTOF System

- High mass accuracy with exceptional negative ion performance for oligonucleotide analysis
- TurboV™ source and QJet® ion guide provide efficient ionization and transmission of analytes while reducing contamination from use of ion-pairing reagents
- Instrument control using SCIEX OS enables accessibility and ease of use regardless of expertise level
- High resolution platform for qualitative and quantitative oligonucleotide workflows
- Integration of data with ProMass for high throughput and robust mass confirmation and purity assessment

## Methods

**Sample Preparation:** Synthetic oligonucleotides were purchased from IDT as custom sequences. The samples were reconstituted in water prior to use.

**Chromatography:** The separation was accomplished using an Agilent 1290 Infinity UHPLC fitted with a BEH Column, 2.1 mm x 100 mm with 1.7  $\mu$ m particles. Mobile phase A was 15 mM diisopropylethylamine (DIEA) with 100 mM 1,1,1,3,3,3-hexafluoro isopropanol (HFIP) in water while mobile phase B was 15 mM DIEA with 100 mM HFIP in 50:50 v/v % methanol:water. The gradient used is shown in Table 1. The column temperature was held at 60°C. An injection volume of 1  $\mu$ L was employed, resulting in 50 ng on column.

**Mass Spec:** A SCIEX X500B QTOF with a Turbo V™ source and TwinSpray probe was used for analysis. The MS parameters are listed in Table 1. The data was acquired using SCIEX OS Software 1.5

**Data processing:** All data were processed with either SCIEX OS or the ProMassProcessor and ProMass for SCIEX software.

## Results and Discussion

The analysis of oligonucleotides can be challenging for a wide variety of reasons. A specific challenge centers around the use of ion-pairing reagents which enable the separation of oligonucleotide species based on order of their charge which is generally indicative of their length. Typically, ion-pairing systems are present at low to mid millimolar concentrations and efficient removal of these components during the electrospray process is critical. In addition, the presence of alkali metal ions, namely sodium and potassium, can complicate analysis. Proper maintenance and regular mobile phase preparation with high quality reagents is required.

As shown in Figure 2, a high quality chromatographic separation is accomplished using ion-pairing reverse phase chromatography. In the work presented here a generic gradient was used which enables analysis of a wide range of oligonucleotide sequences with little requirement for the adjustment of the method.

Averaging the spectra under the main peak results in a typical raw spectrum for oligonucleotides. In this case, charge states spanning from -3 to -10 are shown in Figure 3. As evidenced in the spectrum, adduct formation is quite low and adduct formation is most pronounced in the lowest charge state species. As expected, each charge state demonstrates isotopic resolution. The inset of Figure 3 shows the zoomed view of the -3 charge state for the oligonucleotide

**Table 1. LC Gradient**

Time (min)	%A	%B	Flow Rate ( $\mu$ L min <sup>-1</sup> )
Initial	70	30	300
1.0	70	30	300
6.0	10	90	300
7.0	10	90	300
8.0	70	30	300
10.0	70	30	300

**Table 2. MS Parameters**

	Experiment 1
Polarity	Negative
Intact Protein Mode	On
GS1 (psi)	70
GS2 (psi)	70
Curtain Gas	35
Temperature (°C)	350
TOF Start Mass (Da)	600
TOF Stop Mass (Da)	2900
Accumulation Time (s)	0.25
Spray Voltage (V)	-4500
Declustering Potential (V)	-80
Collision Energy (V)	-10
Time bins to sum	4

Frequently the objective of intact mass measurement is confirmation of the target mass for a specific oligonucleotide sequence. In addition, there may be a need to identify failure sequences or impurities which are present in a sample. A common tool for this processing is ProMass which enables batch processing of oligonucleotide data. To enable the migration of data collected from SCIEX mass spectrometers the ProMassProcessor was used to define the peak selection criteria and sample information for export to ProMass.

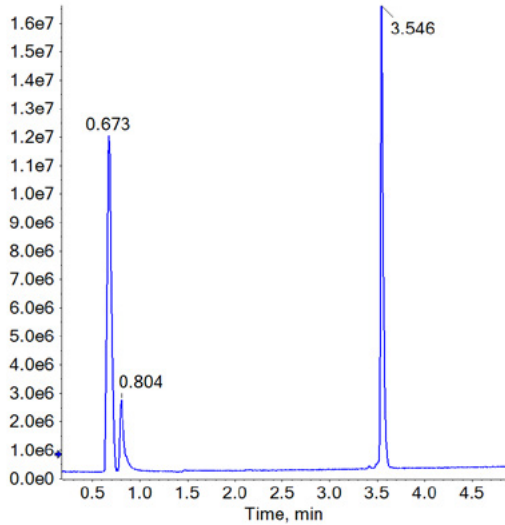


Figure 2. Total Ion Chromatogram for the analysis of a 22-mer oligonucleotide CAT GGT CCT GCT GGA GTT CGT G.

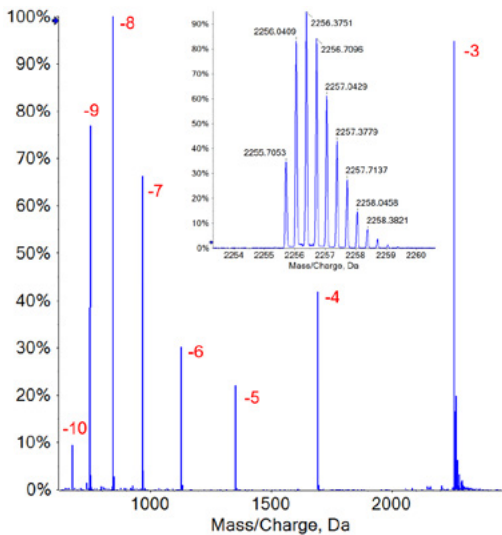


Figure 3. Raw mass spectrum of a 22-mer oligonucleotide CAT GGT CCT GCT GGA GGT CGT G

Within the ProMassProcessor, data files are selected for processing. Files may be move individually or in groups. There is no limit on the number of files that may be moved, however downstream processing in ProMass will increase as more files are selected.

Following file selection, peak selection criteria can be defined for both MS and, if collected simultaneously, the UV data for each experiment as shown in Figure 5A. In addition, peak finding criteria may be defined and additional peak filtering criteria such as minimum and maximum retention time, relative peak area and the maximum number of peaks to find are defined, Figure 5B.

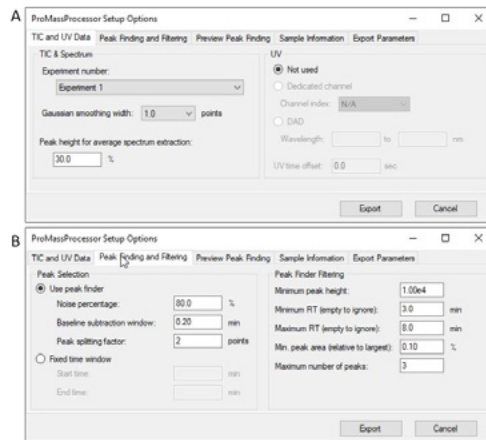


Figure 5. A. Definition of peak selection and experimental criteria in ProMassProcessor. B. Definition of peak finding in ProMassProcessor.

Prior to export it is possible to review the integration of each selected file using the defined criteria, Figure 6. If the integration is not suitable the parameters may be adjusted and will apply to each file. Alternatively, if there is a desire to include a specific peak from a file that falls outside of the defined criteria, or alternatively if a peak should be excluded it is possible to manually add or remove these peaks as shown in Figure 6.

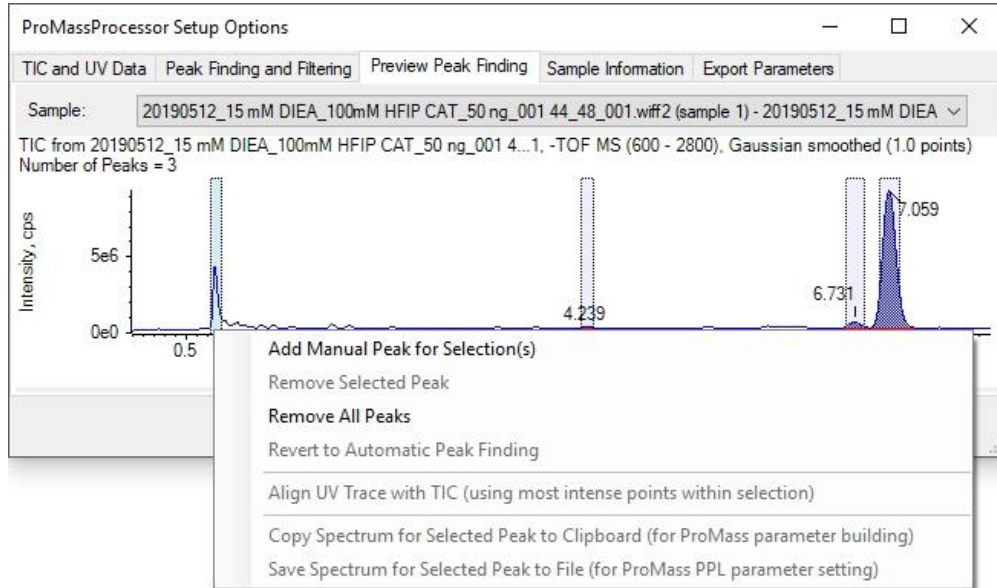


Figure 6. Preview of peak finding results in ProMassProcessor. Additional peaks may be defined manually if desired or already selected peaks may be removed.

For each sample the target information may be defined separately. The target information may be entered as a sequence which is defined using the standard nomenclature supported by ProMass, as a target mass or a combination of both. If more than one target is present in a sample, each species is separated using a comma. To expedite entry, values may be pasted from other sources such as Excel. The software is able to use defined oligonucleotide sequence information, calculate the theoretical mass sequence and, if selected, the associated failure sequences from both the 5' and 3' ends during ProMass processing.

The final stage in processing data is to export the data from the ProMassProcessor to ProMass. In this case the user selects the parameter file defined within ProMass that they would like to apply and the processing begin immediately. To initiate processing a licensed version of ProMass for SCIEX is required.

Following processing the results are displayed accessed through ProMass Browser from the main screen of ProMass, Figure 8. Users can either open the last processed results using the browser button or previously processed results by opening them directly. The results displayed in ProMass provide an easy to understand indication of which samples have passed or failed. The underlying data for each sample is easily investigated by selecting each well and reviewing the total ion chromatogram, raw MS data, and reconstructed MS data. In addition, a table of all identified components with their corresponding abundance, mass error and the sequence the mass corresponds to is also provided.

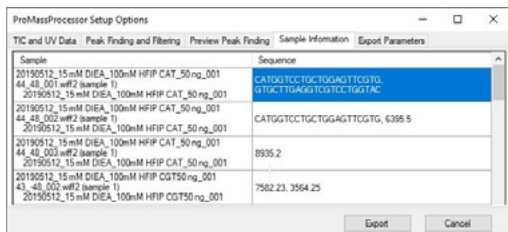


Figure 7. Definition of Sequence and/or mass information in ProMassProcessor.

### Conclusions

- The X500B QTOF system generates high quality oligonucleotide data in the presence of ion-pairing mobile phase components
- High mass accuracy and resolving capabilities enable unambiguous confirmation of oligonucleotide mass
- Low adduct formation is obtainable with suitable cleaning protocols for system components and use of high quality reagents
- Data processing using ProMassProcessor and ProMass provides a semi-automated solution for oligonucleotide characterization and QC

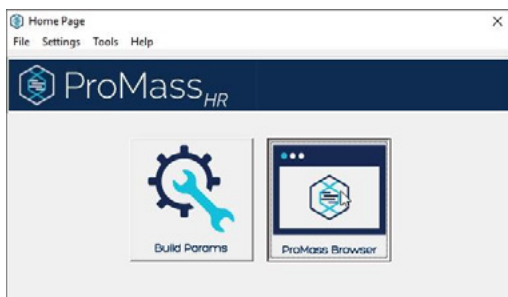


Figure 8. Home page of ProMass software



ProMass Sample Browser												
	1	2	3	4	5	6	7	8	9	10	11	12
A	●	●	●	●	●	●	●	●	●	●	●	●
B	●	●	●	●	●	●	●	●	●	●	●	●
C	●	●	●	●	●	●	○	○	○	○	○	○
D	○	○	○	○	○	○	○	○	○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○	○	○	○	○	○	○	○	○
H	○	○	○	○	○	○	○	○	○	○	○	○

Figure 9. Sample summary from ProMass Sample Browser. Colored wells indicate passing, failing, or marginal results based on criteria defined in ProMass parameters.

RT (min)	Calculated Mass (Da)	Observed Mass (Da)	Mass Error	Intensity	Sequence
9.864	6146.0	6146.6	-1.4 Da (-0.023 %)	1.84E+004	A1-T20 +3Phos
9.864	6221.1	6219.6	-1.5 Da (-0.024 %)	1.92E+004	6372.2 (G depurination)
9.864	6237.1	6235.1	-2.0 Da (-0.032 %)	4.03E+004	6372.2 (A depurination)
9.864	6372.2	6371.1	-1.1 Da (-0.017 %)	1.23E+006	A1-T21
9.496	6043.0	6041.8	-1.2 Da (-0.020 %)	1.79E+003	6372.2 (Minus G)
9.496	6139.0	6138.0	-1.0 Da (-0.016 %)	3.23E+003	C2-T21 +5Phos
9.327	6083.0	6082.1	-0.9 Da (-0.015 %)	8.11E+003	6372.2 (Minus C)
9.048	5849.8	5849.1	-0.7 Da (-0.012 %)	3.69E+003	G3-T21 +5Phos
9.048	6068.0	6067.2	-0.8 Da (-0.013 %)	1.77E+004	A1-T20, 6372.2 (Minus T)
8.571	5520.6	5519.5	-1.1 Da (-0.020 %)	2.05E+003	G4-T21 +5Phos
8.571	6059.0	6058.6	-0.4 Da (-0.007 %)	1.70E+004	C2-T21, 6372.2 (Minus A)
8.291	5769.8	5768.9	-0.9 Da (-0.016 %)	2.02E+004	G3-T21

Figure 10. Individual sample summary table displaying retention time, calculated and observed mass, mass error, intensity, and sequence match based on mass.

AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-10050-A



**Headquarters**  
 500 Old Connecticut Path | Framingham, MA 01701 USA  
 Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
 For our office locations please call the division headquarters or refer to our website at [sciex.com/offices](http://sciex.com/offices)

Download this technical note





# Extending the Lower Limits of Quantification of a Therapeutic Oligonucleotide Through Microflow LC-MS/MS

Featuring **SCIEX QTRAP® 6500+ LC-MS/MS System with OptiFlow™ Turbo V source and the SCIEX M5 MicroLC System**

Daniel Warren, Sean McCarthy, Lei Xiong  
SCIEX, Framingham, MA, USA

## Introduction

Over the past 30 years oligonucleotide therapeutics have occupied a significant space in clinical drug development pipelines offering the exciting potential to target and modulate gene expression. Antisense RNA strands were initially identified as candidates for therapeutic promise in 1998. Fomivirsen, a phosphorothioate antisense oligonucleotide, received the first FDA approval as an oligonucleotide therapeutic. Since this milestone, antisense oligonucleotides still play an important role in modern drug development strategies. The development of ion-pairing reversed phase liquid chromatography (IP-RP LC) methodologies have paved the way for mass spectrometers to become a viable platform for oligonucleotide analysis by providing a strategy for sample introduction and orthogonal front-end separation that is compatible with electrospray ionization techniques common to modern mass spectrometers. IP-RP LC-MS is often the preferred solution for both qualitative and quantitative oligonucleotide analysis; however, several fundamental challenges remain with this approach. The concentrations of ion pairing reagents needed to adequately retain and separate polar oligonucleotides on a reversed phase column, often an alkylamine paired with a fluorinated alcohol, contribute to charge competition resulting in electrospray ion suppression. This phenomenon compromises the mass spectrometer response and may prevent desired limits of detection and quantification from being achieved. The routine use of ion pairing reagents pose another challenge as their accumulation within the mass spectrometer contributes to contamination which accelerates front end cleaning and maintenance intervals. Here a solution is presented to these widely recognized challenges through a novel microflow LC-MS strategy. The sensitivity improvements that result from reduced flow rates entering the MS and the associated enhancement in ionization efficiency have been well documented and prompted microflow LC-MS assays to be deployed for myriad applications. In this instance the reduction of contaminants entering the mass spectrometer is proportional to the mobile phase flow offering an additional advantage of microflow LC-MS. This study was



designed to characterize the improvement of quantification for a phosphorothioate antisense oligonucleotide assay scaled down to microflow LC-MS.

## Key Features of the SCIEX M-Series and OptiFlow™ TurboV Ion source

- M5 MicroLC system provides:
  - Microfluidic flow control for accurate flow rates down to 1  $\mu\text{L}/\text{min}$
  - Direct Inject and trap-elute capabilities for fast and large volume sample loading
  - Flexibility to couple with any microLC column
  - Significantly reduced solvent consumptions and system contamination from ion pairing reagents
- OptiFlow™ Turbo V Source on the QTRAP 6500+ LC-MS/MS system provides:
  - Easy setup with no probe or electrode position optimization required
  - Robust performance and long electrode lifetime

## Materials and Methods

**Samples and Reagents:** All materials were purchased from Sigma Aldrich including HFIP (≥99.8%), DIEA (99.5%), Acetonitrile, Ethylenediaminetetraacetic acid (EDTA) and a custom oligonucleotide standard designed to be structurally equivalent to fomivirsen. Fomivirsen is a synthetic antisense oligonucleotide comprising of 21 bases (5'-G\*C\*G\*T\*T\*T\*G\*C\*T\*C\*T\*T\*C\*T\*T\*C\*T\*T\*G\*C\*G-3') linked by phosphorothioate bonds. Despite its molar mass of 6,682.4 g/mol, fomivirsen can be detected within a mass/charge range suitable for quadrupole analysis as a multicharged species in negative mode electrospray ionization (ESI).

**Chromatography:** A mobile phase system of 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) and N,N-Diisopropylethylamine (DIEA) was chosen to achieve adequate chromatographic retention and resolution when paired with a reversed phase C18 column. These mobile phases are also favored due to compatibility and enhanced ionization efficiency when coupled to MS. The formation of alkali metal adducts are an intrinsic challenge with an LC-MS oligonucleotide analysis as the negatively charged phosphodiester backbone electrostatically attracts positively charged alkali metal salt cations often ubiquitous to HPLC systems and columns. The formation of alkali metal adducts, most commonly but not exclusively Na<sup>+</sup> and K<sup>+</sup> cations, distribute the peak area across an envelope of adduct peaks compromising the targeted oligonucleotide peak intensity. The extent of passivation and materials present in the HPLC system are significant contributors to alkali metal adduct ratios and may bias a performance comparison between two unique HPLC systems, adduct formation was assessed for both conventional flow and microflow systems to ensure less than 10% adduct formation for both flow regimes. Both conventional 2.1mm ID and microbore 0.3 mm ID columns were purchased from Phenomenex, the stationary phases selected primarily on pH and temperature stability. A set of calibrators was prepared from 0.5 - 1000 ng/mL in a solution of 100µM EDTA in mobile phase A. An oligonucleotide internal standard was added to each calibrator at a concentration of 500 ng/mL.

**Table 1. MS Conditions Common to Both Analysis.**

Name	Q1	Q3	DP	CE	CXP	EP
Fomivirsen [M-10H] -10 MRM-1	667	319	-80	-35	-12	-10
Fomivirsen [M-10H] -10 MRM-2	667	304	-80	-45	-12	-10
Fomivirsen [M-9H] -9 MRM-1	741	319	-80	-40	-12	-10
Fomivirsen [M-9H] -9 MRM-2	741	344	-80	-40	-12	-10
Fomivirsen [M-8H] -8 MRM-1	834	319	-80	-45	-12	-10
Fomivirsen [M-8H] -8 MRM-2	834	649	-80	-35	-12	-10

**LC-MS Conditions for Comparative Analysis:** To identify the sensitivity difference between the LC-MS and microflow LC-MS analysis, each sample was analyzed in triplicate on the same QTRAP 6500+ LC-MS/MS system without modification to the instrument parameter settings.

As the observed charge state for oligonucleotides may be influenced by the reversed phase (RP-IP) mobile phase system and LC-MS conditions, charge state characterization and MRM method development were performed by injecting standards on column rather than a more conventional infusion approach. Low molecular weight fragment ions (< 200 Da m/z) such as phosphate ions were not selected during MRM method development due to poor specificity despite often offering superior intensity. The MRM parameters described in Table 1 were kept consistent between the two methodologies while Tables 2 and 3 contain the source and liquid chromatography settings that were optimized for each specific hardware configuration and flow rate. First, the mass spectrometer was coupled to an Exion™ UHPLC system and Ion Drive Turbo V™ ion source to benchmark the LC-MS performance then outfitted with a SCIEX M-series µUHPLC system and an OptiFlow™ Turbo V Source with a 25 µm SteadySpray™ probe and electrode for microflow LC-MS analysis. To attain a consistent linear velocity between the two experiments, the flow rate used for the microflow was simply scaled proportionately to the reduced column volume, the flow profile was slightly adjusted to account for the unique C18 stationary phases.

## Assessment of Chromatographic Performance

Chromatographic resolution and peak shape play a critical role in the performance of any quantitative LC-MS assay. Peak symmetry and width are important factors that contribute to analyte signal to noise (S/N) and therefore influence the sensitivity and lower limits of quantification (LLOQ) of an assay. To properly evaluate the improved ionization efficiency at low microliter flow rates, it was essential to design microflow methodology with comparable chromatography to that of common conventional flow UHPLC results. In the absence of an LC system designed specifically for microflow separations, excessive dead volume and extra column peak band broadening frequently results in decreased throughput and poor peak shapes which prevent the benefits of microflow LC-MS from being realized. The low dead volume and precise solvent delivery of the SCIEX M-series system permitted throughput and chromatographic resolution to be equivalent to that typical of LC-MS result performed above 200 µL/minute (Figures 1 and 2).

Table 2. Conventional LC-MS Method Parameters.

Source/Gas Parameter	Value	Source/Gas Parameter	Value
Curtain gas:	30	CAD gas:	High
Ion source gas 1:	70	Ion spray voltage:	-4500
Ion source gas 2:	70	Source temperature:	500

Parameter	Value
Stationary phase	Phenomenex Clarity® Oligo-MS 50 x 2.1 mm; 2.6µm particle; 100Å pore
Mobile phase A	15mM DIEA 100mM HFIP in water
Mobile phase B	15mM DIEA 100mM HFIP in 50% methanol
Flow rate	300 µL/min (3.6 mm/sec linear velocity)
Column temperature	60°
Injection volume	5 µL

Time	Flow Rate (µL/min)	%A	%B
0	300	70	30
0.5	300	70	30
3.5	300	10	90
4.0	300	10	90
4.1	300	70	30
5.0	300	70	30

Table 3. Microflow LC-MS Method Parameters.

Source/Gas Parameter	Value	Source/Gas Parameter	Value
Curtain gas:	25	CAD gas:	High
Ion source gas 1:	40	Ion spray voltage:	-4500
Ion source gas 2:	40	Source temperature:	350

Parameter	Value
Stationary phase	Phenomenex Gemini® C18 50 x 0.3 mm; 3µm particle; 100Å pore
Mobile phase A	15mM DIEA 100mM HFIP in water
Mobile phase B	15mM DIEA 100mM HFIP in 50% methanol
Flow rate	6 µL/min (3.5 mm/sec linear velocity)
Column temperature	60°
Injection volume	5 µL

Time	Flow Rate (µL/min)	%A	%B
0	6	80	20
0.5	6	80	20
3.0	6	10	90
4.0	6	10	90
4.1	6	80	20
5.0	6	80	20

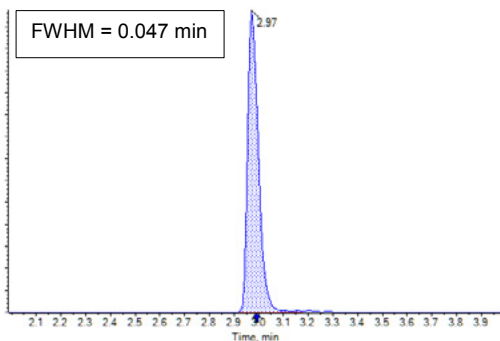


Figure 1. Example Chromatography Performance Achieved for Fomivirsen at 300 µL/minute.

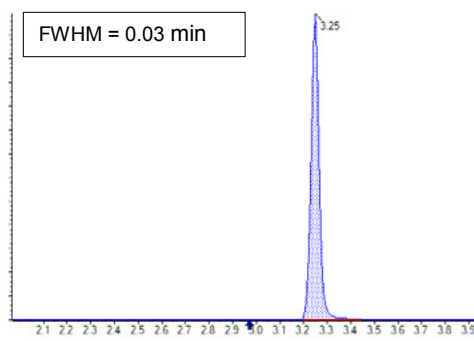


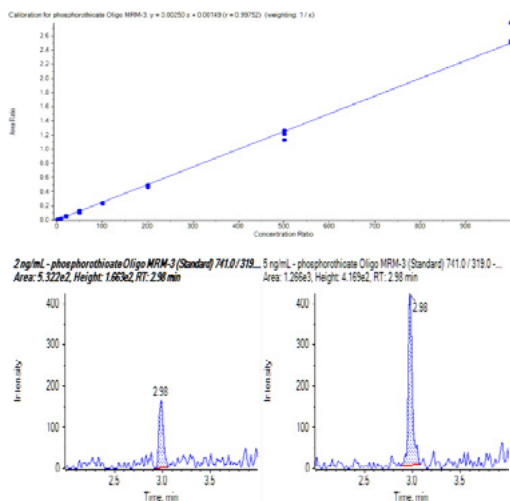
Figure 2. Example Chromatography Performance Achieved for Fomivirsen at 6 µL/minute.

[Download this technical note](#)

## Quantitation Results

Acceptance criteria of  $\pm 15\%$  accuracy and CV of  $< 20\%$  ( $< 20\%$  at LLOQ) was used to assess the linear range of quantification for both LC-MS and microflow LC-MS experiments. A linear regression with  $1/x$  weighting was applied to the peak areas for both data sets. Data processing was performed using MultiQuant Software 3.0. A linear concentration range from 5 -1000 ng/mL was achieved on the conventional LC-MS configuration with an LOD of 2 ng/mL. Figure 3 describes the LC-MS calibration curve results and provides representative LOD and LLOQ peaks.

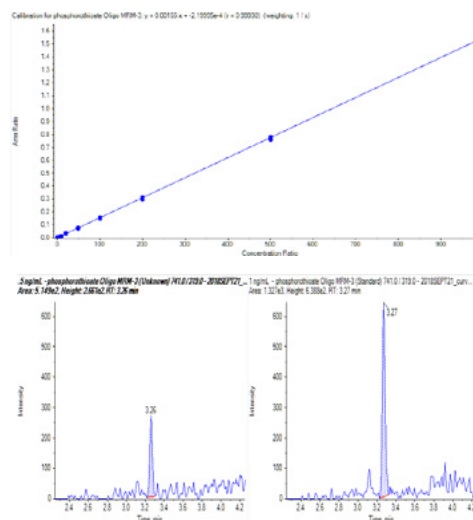
The set of calibrators prepared from 0.5 - 1,000 ng/mL was re-acquired on the QTRAP 6500+ system coupled with microLC. The presented microflow assay for fomivirsen demonstrated the significant improvement upon the LLOQ of 5 ng/mL achieved using the conventional flow LC-MS configuration. Using the same processing parameters and quantitative MRM transition, an LLOQ of 1 ng/mL was attained with an LOD of 0.5 ng/mL. The linear range of quantification from 1- 1,000 ng/mL. Figure 4 depicts the LOD and LLOQ XICs attained through microflow as well as the calibration curve results.



**Figure 3. Quantitation Results for the Conventional Flow Experiment (300  $\mu\text{L}/\text{min}$ ).** (Top) Calibration Curve for fomivirsen with  $r = 0.99752$  and weighting of  $1/x$ . (Bottom) Extracted ion chromatograms (XICs) of selected MRM for fomivirsen for the LOD of 2 ng/mL (left) and the LLOQ of 5 ng/mL (right).

Table 4 provides a full summary of the quantitative performance of the microflow LC-MS configuration.

To further assess the improved ionization efficiency and the associated improvements in peak area and signal to noise, Figure 5 offers a comparison of peak intensity, area, and signal to noise for the LLOQ and ULOQ achieved at conventional flow rates, 2 ng/mL and 1,000 ng/mL. Figure 6 plots a comparison of the internal standard peak area across the duration of each data sample set.



**Figure 4. Quantitation Results for the Microflow Experiment (6  $\mu\text{L}/\text{min}$ ).** (Top) Calibration Curve for fomivirsen with  $r = 0.99930$  and weighting of  $1/x$ . (Bottom) Extracted ion chromatograms (XICs) of selected MRM for fomivirsen for the LOD of 0.52 ng/mL (left) and the LLOQ of 1 ng/mL (right).

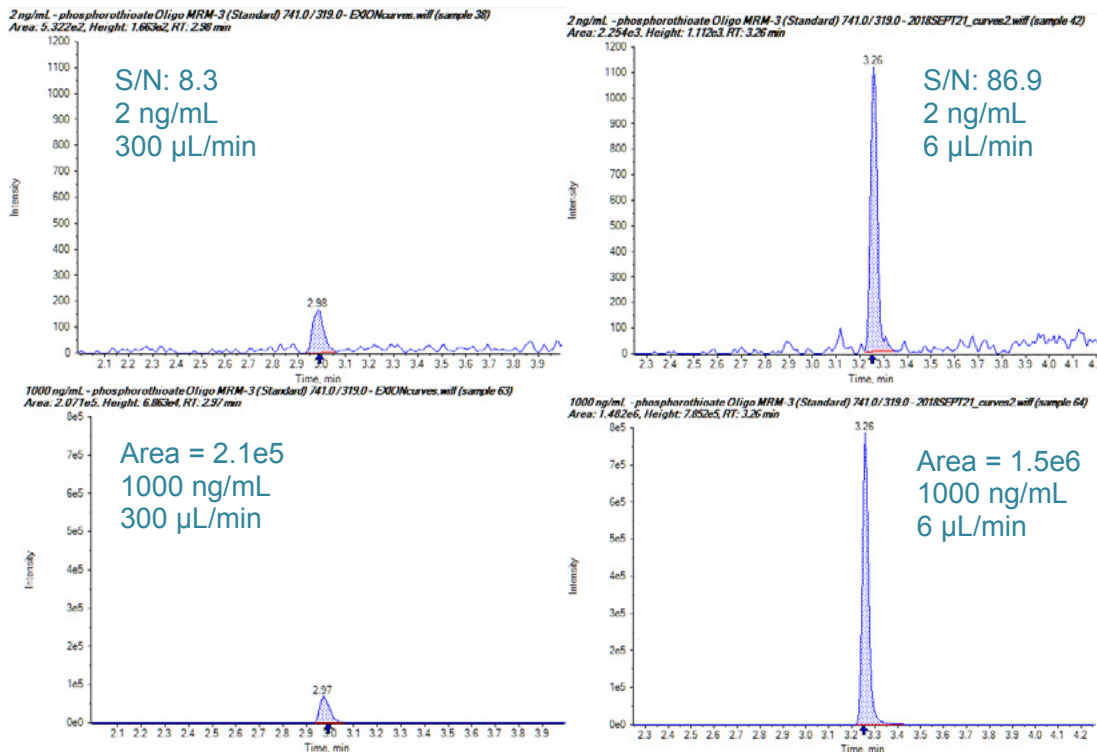


Figure 5. Extracted Ion Chromatograms (XICs) of Selected MRM Transitions. Data is shown for fomivirsen using conventional LC-MS (left) and microflow LC-MS (right) at 2 ng/mL (top) and 1000 ng/mL (bottom). Signal to noise calculations performed using MultiQuant software.

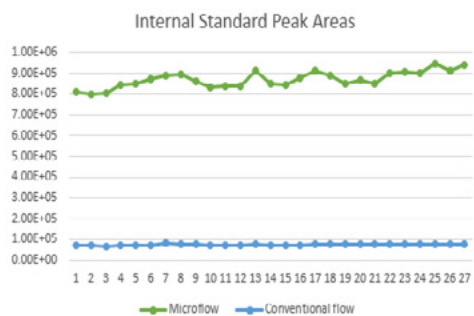


Figure 6. Metric Plot of Oligonucleotide Internal Standard Peak Areas for Microflow (green) and Conventional Flow (blue) Analysis.

Table 4. Microflow LC-MS Calibration Curve for fomivirsen at 6 µL/minute;  $r = 0.99930$  (weighting:  $1/x$ ).

Actual Conc. (ng/mL)	Calculated Conc. (ng/mL)	Accuracy (%)	CV (%)
0.5	0.62	123.80	16.07
1	1.02	101.87	14.87
2	1.92	96.19	0.89
5	4.59	91.78	8.96
10	9.06	90.64	6.81
20	20.61	103.04	3.65
50	47.96	95.91	3.94
100	98.47	98.47	2.85
200	196.08	98.04	2.91
500	494.46	98.89	1.70
1000	1013.72	101.37	4.82

### Conclusions

MRM based LC-MS and microflow LC-MS assays were developed to assess the quantitative performance of a therapeutic oligonucleotide acquired at varied HPLC flow rates. The results presented indicate a significant improvement in peak area, peak intensity, and signal to noise when a dedicated microflow HPLC system running at 6  $\mu\text{L}/\text{minute}$  is utilized compared to a conventional LC-MS system running 300  $\mu\text{L}/\text{minute}$ . In addition to the improvement in sensitivity and overall quantitative performance, it is also notable that the microflow LC-MS assay dramatically reduced the required mobile phase additives that inherently contribute to front end contamination of the mass spectrometer.

AB Sciex is doing business as SCIEX.

© 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-8916-A

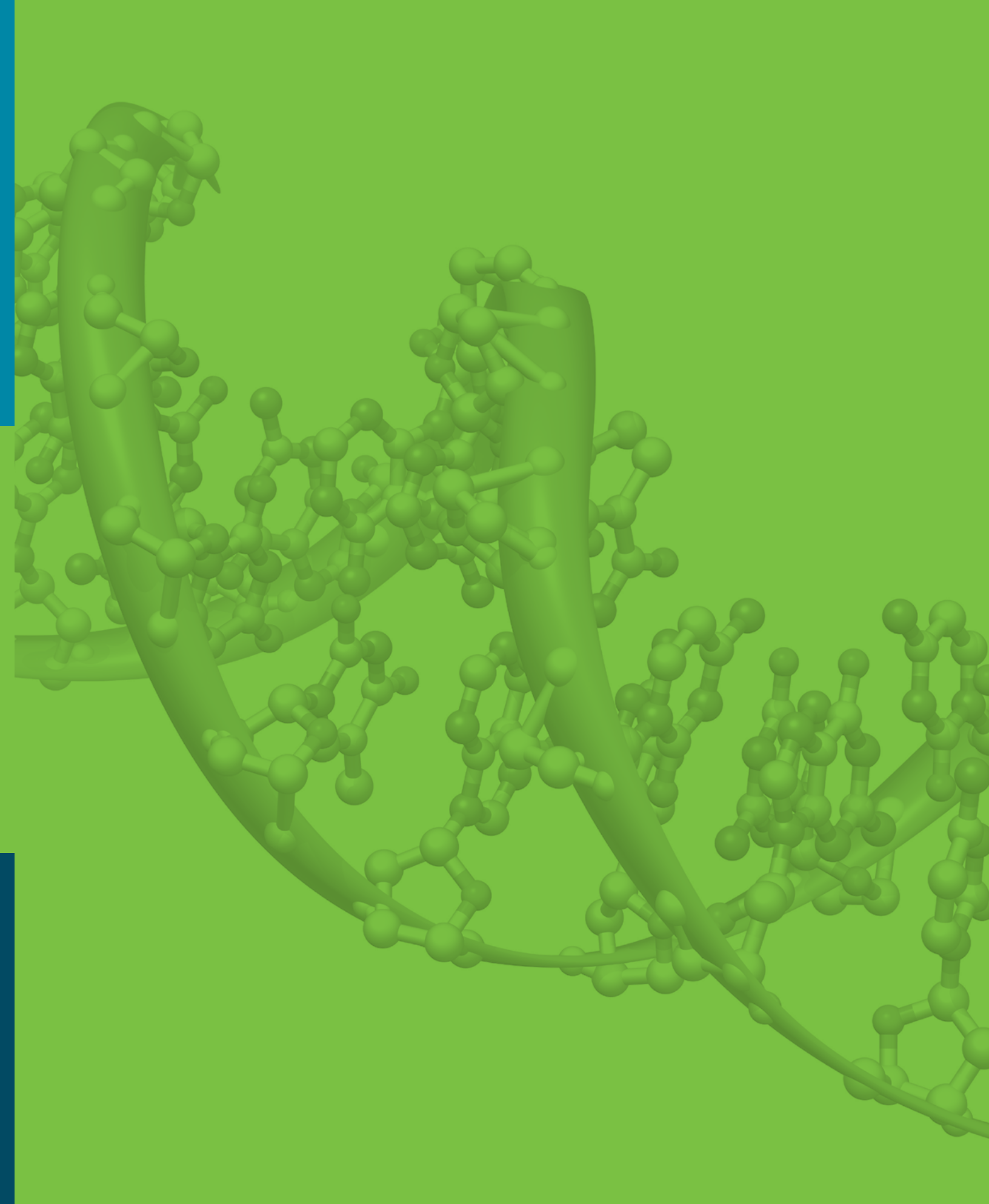


**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

[Download this technical note](#)





# Quantification of Large Oligonucleotides using High Resolution MS/MS on the TripleTOF® 5600 System

Thomas Knapman<sup>1</sup>, Vicki Gallant<sup>1</sup> and Martyn Hemsley<sup>2</sup>

<sup>1</sup>SCIEX, Warrington, UK, <sup>2</sup>Covance, Harrogate, UK

## Key Challenges of Oligonucleotide Bioanalytical Assay

1. The bioanalysis of oligonucleotides as therapeutics requires sensitive, specific and robust analysis.
2. Many ELISA-based oligonucleotide measurements do not accurately distinguish large metabolites from the full-length oligonucleotide of interest.
3. ELISA- and UV-based measurements have limited dynamic range which complicates quantitative analysis of oligonucleotides in complex matrices.

## Key Benefits of MRM<sup>HR</sup> Workflow for Oligonucleotide Bioanalytical Assay

1. High sensitivity MS/MS enables the quantitative MRM<sup>HR</sup> workflow, providing high selectivity in biological matrices.
2. High resolution, accurate mass MS/MS spectra enable qualitative verification of oligonucleotide sequences.
3. The MRM<sup>HR</sup> workflow offers a dynamic range of two to three orders of magnitude.

## Unique Features of MRM<sup>HR</sup> Workflow on TripleTOF® 5600 System

1. Summing of multiple ion transitions to increase both sensitivity and selectivity of quantitation.
2. Accelerated method development times, since ion transitions can be selected post-acquisition to eliminate background interferences.
3. High multiplexing due to high acquisition speeds (up to 100 spectra per second) for simultaneous quantitation of multiple species, including multiple oligonucleotide sequences and/or their metabolites.



Figure 1: The TripleTOF® 5600 System has the speed and sensitivity to deliver high-throughput targeted quantitation of many species in a single run. This example focuses on using MRM<sup>HR</sup> workflow to quantify a synthetic oligonucleotide from human plasma, demonstrating post-acquisition fragment ion selection and summing product ions to achieve the highest possible sensitivity and selectivity.

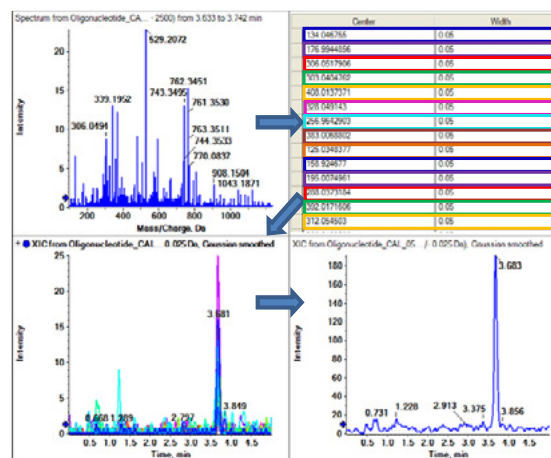


Figure 2: Principle of MRM<sup>HR</sup> Workflow Quantitation. Looped full scan MS/MS spectra are acquired for each precursor. Selected fragments are then extracted post-acquisition using narrow extraction widths (in this case 50 mDa) to produce high resolution XICs. These XICs can be used individually or summed, depending on which provides the best selectivity and sensitivity for quantitation.

## Introduction

Quantitative analysis of synthetic oligonucleotides in biological matrices is an important aspect of pharmacokinetic (PK), toxicokinetic (TK) and metabolic pathway studies in drug development<sup>1</sup>. With an increasing number of oligonucleotide based drugs in research pipelines, the acceleration of the drug development process by reducing the time spent on method development, and by performing simultaneous qualitative structural analysis with quantitative analysis are crucial advantages in any potential quantitation approach<sup>1</sup>.

Current LC-MS approaches to oligonucleotide quantitation predominantly use multiple reaction monitoring (MRM), however the complex fragmentation pathways of oligonucleotide species coupled with the variability of matrix effects mean that it can be difficult to predict the sensitivity and selectivity of a given MRM transition without significant optimization<sup>1</sup>. These effects limit the utility of low resolution quantitation methods both in terms of the achievable limits of quantitation and in sample throughput, particularly when quantifying large numbers of potential drug candidates of different sequences, and their metabolites.

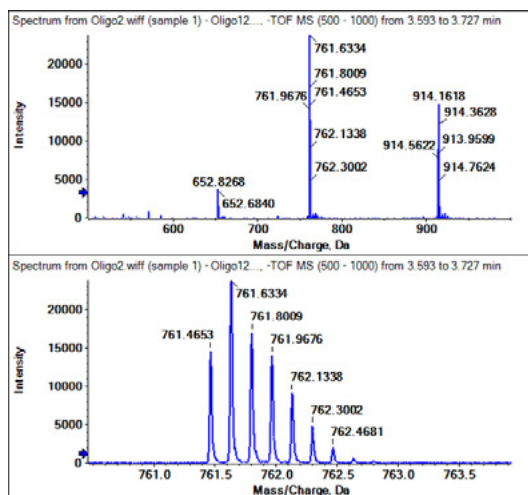


Figure 3: TOFMS Analysis of Intact Oligonucleotide 1 acquired on the TripleTOF<sup>®</sup> 5600 system (top). Isotope resolution of higher charge states is achieved with a resolution of 30,000–40,000, as illustrated by zooming in on the 6<sup>+</sup> charge state (bottom). This resolution is achieved at TOF MS scan speeds of 10 MS spectra per second.

## Materials and Methods

### Sample Preparation

The synthetic DNA Oligonucleotide 1 was spiked into human plasma over a concentration range of 0.025 to 10 nM. Oligonucleotide 2 was used as an internal standard.

### LC Conditions

LC System	Shimadzu Prominence XR UFLC
Analytical column	Waters Acquity BEH, 50 x 2.1 mm, 1.7 $\mu$ m, temp.= 60 °C
Analytical flow	0.40 ml/min (initial 24 hour flush)
Mobile Phase A	Water (15 mM TEA, 400 mM HFIP)
Mobile Phase B	50:50 Methanol:Water (15 mM TEA, 400 mM HFIP)

### Gradient conditions

Time (min)	Mobile phase A%	Mobile phase B%
0.5	75	25
5	50	50
5.1	0	100
5.5	0	100
5.6	75	25
6	75	25

### MS Conditions

MS System	TripleTOF <sup>®</sup> 5600+ system with a DuoSpray <sup>™</sup> Source
Ionization Mode	ESI with Negative Mode
TOF MS range	m/z 100–2500 at 250 msec accumulation time
MRM <sup>HR</sup>	2 product ions each 250 msec
Collision energy spread	-40 $\pm$ 4 eV
Source temperature	550 °C

### Software

Data acquisition	Analyst TF <sup>™</sup> 1.5.1 Software
Data review	PeakView <sup>™</sup> 1.2 Software
Deconvolution	BioAnalyst <sup>™</sup> Software
Quantitation	MultiQuant <sup>™</sup> 2.1 Software

## Results and Discussion

**MS and MS/MS Analysis of Oligonucleotides.** TOFMS analysis of Oligonucleotide 1 showed a charge state envelope consisting of  $[M-5H]^{5-}$ ,  $[M-6H]^{6-}$  and  $[M-7H]^{7-}$  ions (Figure 3) with a resolution of approximately 36,000. Inspection of the TOFMS spectrum showed that the system passivation process had reduced adduct formation to less than 5% relative to the fully protonated form, thus facilitating quantitation from the  $[M-6H]^{6-}$  peak (data not shown).

The principle of the MRM<sup>HR</sup> workflow for quantitation is to acquire full scan TOF MS/MS spectra for each species of interest, and to use high resolution extracted ion chromatograms (XICs) for quantitation, summing multiple transitions where appropriate to achieve optimum sensitivity and selectivity (Figure 2). To develop an MRM<sup>HR</sup> workflow assay for oligonucleotides 1 and 2, full scan MS/MS spectra were acquired for m/z 761.9 for Oligonucleotide 1, and m/z 745.6 for Oligonucleotide 2. The full scan MS/MS spectra were also used to verify the sequence of Oligonucleotide 1 and 2 (Figure 4). The MS/MS spectra were deconvoluted using BioAnalyst™ Software to enable the singly and multiply-charged fragment ions to be plotted on a mass scale. The sequences were subsequently verified by matching theoretical sequence ions to the fragment ions observed in the deconvoluted spectra.

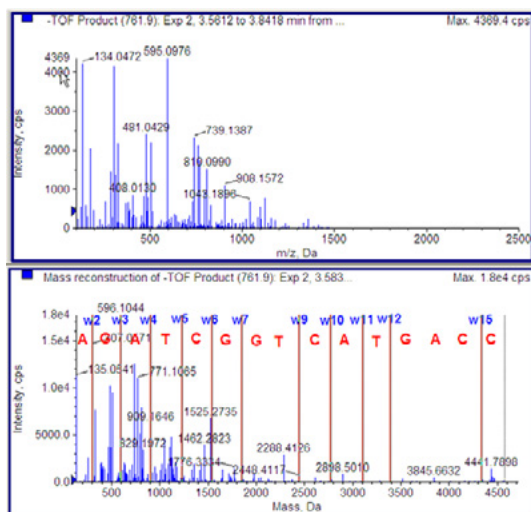


Figure 4: High Quality MS/MS of Oligonucleotide 1 for Characterization and Quantitation Method Development. (top) The full scan MS/MS approach of MRM<sup>HR</sup> workflow enables full qualitative analysis of any targeted analyte. Full sequencing of Oligonucleotide 1 was achieved on the mass reconstructed MS/MS spectrum (bottom) for fragment ion selection.

**MRM<sup>HR</sup> Workflow Assay Development.** Having acquired full scan TOF MS/MS spectra, fragment ions can be selected and extracted post-acquisition for use in quantitation. Figure 2 shows the stepwise process of selecting fragment ions from a full scan MS/MS spectrum, generating multiple high resolution XICs from the selected fragment ions, and summing the XICs to optimize the signal-to-noise. For transitions arising from Oligonucleotides 1 and 2, an extraction width of 50 mDa was used to generate XICs; however it is possible to collect the MS/MS spectra in high resolution mode (>30,000 resolution) and extract with narrower windows to improve selectivity in any given assay if required.

The post-acquisition selection of fragments is a significant advantage of the MRM<sup>HR</sup> workflow, since the selectivity of specific fragment ions in matrix cannot necessarily be predicted prior to data acquisition. In the case of Oligonucleotide 1, the three most intense fragment ions in the MS/MS spectrum gave poor selectivity when extracted (Figure 5), and therefore could not be used for quantitation. In contrast, other less intense transitions showed excellent selectivity, and were subsequently included in the assay. The final XIC trace was achieved by summing 25 different fragment ion XICs (Figure 5). The final data processing using the summed XICs was performed using MultiQuant™ Software.

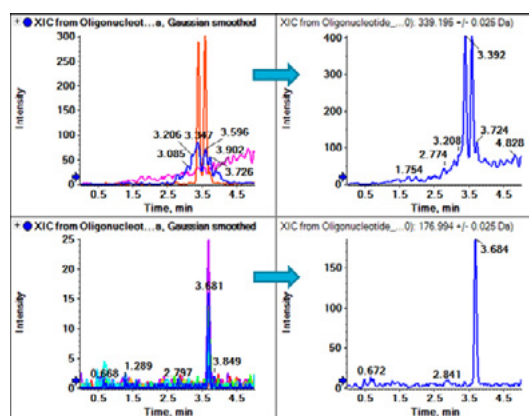


Figure 5: Post-Acquisition Extraction of Structurally Specific Ions. In the case of Oligonucleotide 1, the three most intense fragment ions (top left) are non-selective in plasma at low concentrations and therefore summing of these XICs does not provide a good assay (top right). Because the full scan MS/MS spectrum is acquired in the MRM<sup>HR</sup> workflow, this allows different fragment ions to be selected and extracted for quantitation post-acquisition (bottom left, summed bottom right), and therefore requires significantly less method development than traditional MRM approaches.

**TOF MS vs. MS/MS Quantification Strategies.** The MRM<sup>HR</sup> workflow for quantitation from complex matrices offers significant advantages over the full scan TOF MS approach to quantitation, due to both the high selectivity of the MS/MS based XICs and the ability to remove fragment ions with background interferences, which can impact significantly on the achievable limits of detection and quantification.

Figure 6 shows a comparison of the TOF MS and MRM<sup>HR</sup> workflows applied to the analysis of Oligonucleotide 1. In the case of the TOF MS quantitation, each peak in the isotopic envelope of the [M-6H]<sup>-</sup> charge state was extracted using a 10 mDa extraction window (Figure 6, left). The extracted ion chromatograms from the TOF MS approach show significant matrix interferences that seriously impact the limit of quantification, which is approximately 0.5 nM using this approach.

In contrast, the specificity of the MRM<sup>HR</sup> workflow produced significantly lower limits of detection and quantification (Figure 6, right). This improvement is due to the significant reduction in noise from the complex matrix.

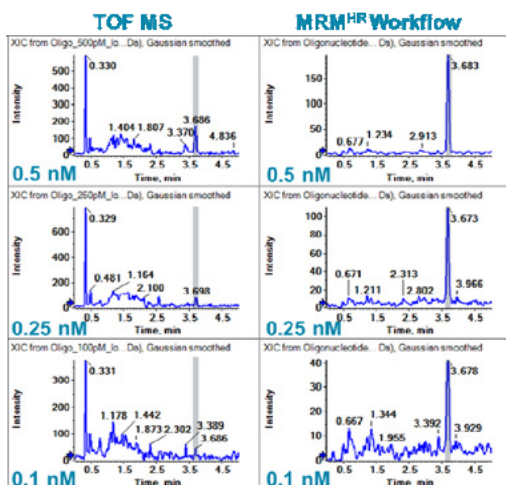


Figure 6: Selectivity of MRM<sup>HR</sup> Workflow in Complex Matrices Allows Better LLOQs to be Obtained. In the case of Oligonucleotide 1, background interferences in full scan TOF MS result in higher limits of detection and quantitation (0.5 nM), while the selectivity of the MRM<sup>HR</sup> workflow allows quantitation of concentrations less than 0.1 nM.

Figure 7 shows the calibration plot of Oligonucleotide 1 using the MRM<sup>HR</sup> workflow with Oligonucleotide 2 used as an internal standard. The %CV values for each concentration analyzed are shown in the embedded table. The correlation coefficient of the response is in excess of 0.99 over ~2 orders of magnitude; the linearity of response at concentrations in excess of 10 nM was not investigated. The lower limit of quantification (LLOQ) using

the MRM<sup>HR</sup> workflow was 0.05 nM (Figure 7), while the LLOQ was 10 fold higher when the same oligonucleotide was analyzed by full scan TOF MS workflow (Figure 6).

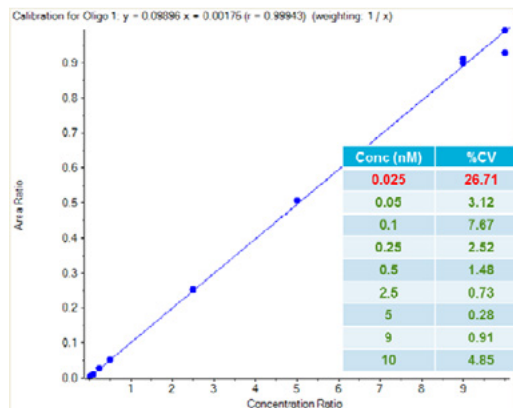


Figure 7: Standard Concentration Curve for Oligonucleotide 1 in Matrix using MRM<sup>HR</sup> Workflow. Concentration curve for Oligonucleotide 1 in plasma, using Oligonucleotide 2 as an internal standard. Using MRM<sup>HR</sup>, excellent linearity was observed, with a lower limit of quantitation of 0.05 nM.

## Conclusions

1. The TripleTOF® 5600 system offers sensitive, high-resolution analysis of large oligonucleotides, with the opportunity to perform both qualitative and quantitative analysis in a single run.
2. Using a targeted MRM<sup>HR</sup> workflow, looped full scan TOF MS/MS spectra of a ~4.5 kDa synthetic oligonucleotide were acquired with a resolution of ~16,000 and XICs were generated from specific fragment ions to achieve a highly sensitive and selective quantitative assay.
3. The ability to quantify oligonucleotides from complex matrices with minimal assay optimization offers the opportunity for high-throughput analysis of potential oligonucleotide-based therapeutics. Upfront method development is highly simplified and consists of specifying a theoretical oligonucleotide m/z and a collision energy, the remainder of the analysis being done post-acquisition.

## References

1. Z. J. Lin, W. Li and G. Dai, J. Pharm. Biomed. Anal. 44



### Considerations for Handling Therapeutic Oligonucleotide Reference Standard and Sample Extraction Tips

Cindy Sanderson  
SCIEX, Framingham, USA

#### Handling Reference Standard Material

The Oligonucleotide API (active pharmaceutical ingredient in GMP Studies) or test article (in GLP Studies) is highly hygroscopic and the water content of an oligonucleotide is related to its environment. Additionally, concentrated solution such as those observed in a clinical setting can be very viscous; both of these traits can lead to inaccuracies in the Assay. To prevent inaccuracies due to the hygroscopic nature of the API, one of two options are presented. The drug substance can be handled only in a room that enables precise temperature and humidity control. The glaring limitation of this approach is that not every site has this available. The alternate option is to equilibrate the API to ambient temperature and humidity prior to handling and using the UV absorbance (typically at 260 nm) and purity to determine the actual concentration. The latter approach assumes knowledge of the oligonucleotide purity. The viscosity issue can be overcome by relying on a gravimetric technique as opposed to a volumetric one.

#### Dealing with Non-Specific Binding

Oligonucleotides are prone to non-specific binding with the container/closure system, components in biological matrices and components of the LCMS system used for analysis. Non-specific binding to sample containers are more pronounced at lower concentrations and can be helped by either storing the solutions in Type I glass containers or using EDTA or other preservative with plastic container. Regardless of oligonucleotide concentration EDTA is a useful preservative as it will chelate divalent cations which are required for nuclease activity. To prevent non-specific binding to components of the LCMS system, one option is to replace as much stainless steel as possible and replace with PEEK or similar tubing; of course this is only applicable to relatively low pressure systems as the PEEK tubing will not tolerate as high a pressure. Another little trick is to add a minute amount ( $\mu\text{M}$  quantities) of EDTA to the mobile phase again to chelate any divalent cations present. Dealing with the non-specific binding of oligonucleotides to components in the biological matrices is more involved and intricate given that oligonucleotides carry lots of negative charge on their phosphodiester-based backbone.

AB Sciex is doing business as SCIEX.

© 2017 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners.

AB SCIEX™ is being used under license. Document number: [RUO-MKT02-6639-A]

The negative charge imparts a strong affinity for ubiquitous cations such as  $\text{Na}^+$  and  $\text{K}^+$  as well as other matrix components.

#### Sample Extraction Tips

There are many ways to separate these matrix components from the therapeutic oligonucleotide including protein precipitation, liquid-liquid extraction, solid phase extraction and various combinations of these procedures. Simple protein precipitations with organic solvents such as acetonitrile are met with limited success as they are prone to low recoveries and the ubiquitous cations are not necessarily removed.

Several liquid-liquid extraction techniques have met with some success, particularly using a phenol/chloroform solution for the LLE. Often, the LLE will include a step that adds a detergent or other modifier to break up any complexes between the oligonucleotide and matrix components. The proteins will partition into the organic layer while the oligonucleotide is left in the aqueous portion along with other polar matrix components. These remaining matrix components can and do interfere with LCMS. For this reason, the liquid-liquid extract can be further treated by solid phase extraction. Adsorption of the oligonucleotide to a reverse phase solid packing material can be enhanced by the addition of a modifier to both the loading and elution solvent; typically an ion pairing agent and/or ammonium hydroxide are used. Once the oligonucleotide is adsorbed onto the solid phase, washing and elution. This combination of LLE followed by SPE is laborious, time consuming and the ability to automate for a larger number of samples is limited. A few years ago Phenomenex came out with their Clarity OTX system which is a mixed-mode (weak anion exchange and reverse phase) SPE cartridge along with buffers designed to work with the cartridge to clean up oligonucleotide samples. The Clarity OTX offers a quicker method for sample preparation from biological matrices with adequate recoveries.



**Headquarters**  
500 Old Connecticut Path | Framingham, MA 01701 USA  
Phone 508-383-7700  
[sciex.com](http://sciex.com)

**International Sales**  
For our office locations please call the division  
headquarters or refer to our website at  
[sciex.com/offices](http://sciex.com/offices)

Download this technical note





The Power of Precision

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. Rx Only.  
Product(s) not available in all countries. For information on availability, please  
contact your local sales representative or refer to <https://sciex.com/diagnostics>.  
All other products are For Research Use Only. Not for use in Diagnostic Procedures.  
© 2020 DH Tech. Dev. Pte. Ltd. RUO-MKT-03-10986-A