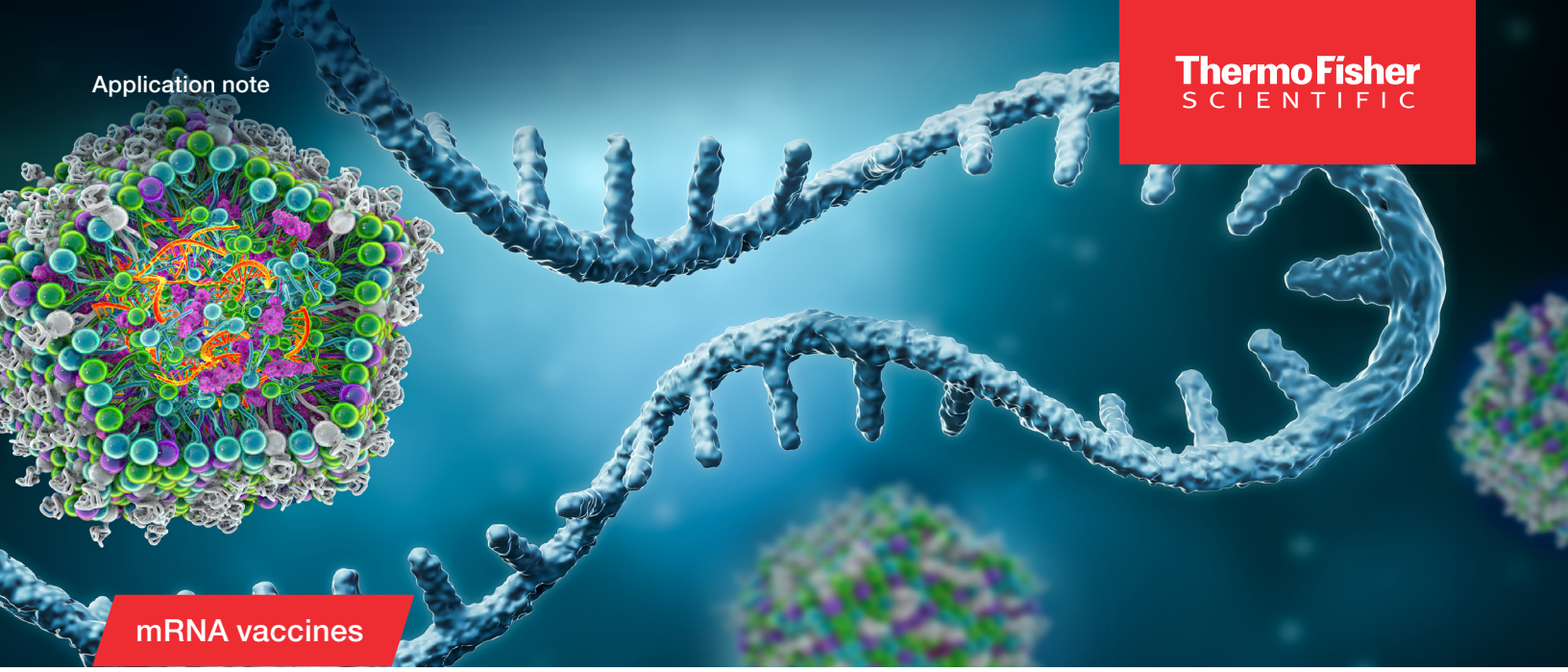


Automated workflow for mRNA *In Vitro* Transcription (IVT)

Keywords
Automation, plasmid purification, sanger sequencing, *in vitro* transcription, mRNA, dPCR, qPCR, residual DNA, QC

In this application note, we describe:

- An automated workflow for *in vitro* transcription (IVT) of mRNA
- Automated plasmid purification and quality control
- Plasmid-directed *in vitro* transcription of RNA on magnetic beads (solid-phase IVT)
- Quality control of purified RNA by direct sequencing and quantification by PCR
- Precise measurement of residual template DNA in purified RNA samples by dPCR

Introduction

The remarkable success of messenger RNA (mRNA) technologies during the COVID-19 pandemic has ushered in a new era of mRNA-based medicines (Pardi et al., 2018; Parhiz et al., 2024). mRNA vaccines and therapeutics utilize mRNA molecules to stimulate the body's immune response or produce therapeutic proteins and demonstrate incredible promise in addressing unmet medical needs in disease prevention and treatment.

One of the advantages of mRNA-based medicines is their versatility and adaptability (Metkar et al., 2024; Qin et al., 2022). Scientists can rapidly develop mRNA sequences to produce specific proteins, allowing for a swift response to emerging diseases or evolving medical needs. This flexibility makes mRNA-based therapeutics a highly attractive option for personalized medicine and tailored treatments.

In vitro transcription (IVT) is the gold standard method for large scale production of mRNA drug substances (Webb et al., 2022). During the discovery phase of new mRNA medicines, the screening process can be quite labor-intensive and challenging. It involves considering various elements such as codon optimization, nucleotide

modifications, UTRs, capping, poly(A) tail, and optimizing reaction conditions (Popova et al., 2024; Rosa et al., 2022; To & Cho, 2021). However, to simplify this process, automating the necessary steps for plasmid purification, *in vitro* transcription and RNA purification can significantly streamline the early discovery phase.

To identify the optimal conditions during the screening process, analytical methods are needed to assess at a minimum, the yield, purity and identity of the DNA template and purified mRNA. Analytical methods for these and additional quality attributes are even more critical for later development phases and require increasing levels of qualification and validation to ensure high quality of the template DNA and mRNA drug substances and products maximizing the therapeutic benefits of mRNA-based treatments (Qin et al., 2022). It is essential that the chosen methods are fit for purpose, phase-appropriate, and adhere to regulatory standards.

Here we demonstrate a complete automated workflow for *in vitro* transcription starting from high-throughput mini-prep scale plasmid purification, through plasmid template linearization to RNA synthesis and purification, as shown in Figure 1. Up to 96 reactions can be set up in parallel on the Thermo Scientific™ KingFisher™ instruments for automated screening. Scale up of *in vitro* transcription and purification of the most promising constructs can be performed on the same instrument in a 24 well format.

To help ensure quality, we developed analytical methods for critical attributes such as concentration, identity, and purity of nucleic acids throughout the workflow (Figure 1): The concentration and purity of plasmid DNA (pDNA) and RNA are measured using UV- absorbance and Qubit per standard methods. The identity of plasmids and RNA sequences are verified by Sanger sequencing. Impurities from the process, i.e. residual *E. coli* DNA in the pDNA preparation and residual pDNA in the purified RNA, are measured by dPCR. Finally, the concentration of the purified RNA is determined by Applied Biosystems™ TaqMan qPCR.

Materials and methods

Automated plasmid DNA purification, IVT and mRNA purification on KingFisher Apex

The VB220707-1149zkx vector for mRNA (pT7[mRNA]-{5utr-FLuc-3utr-pA}) from VectorBuilder Inc. was used for plasmid purification and solid-phase IVT. The plasmid was transformed in VB UltraStable *E. coli* strain. Briefly, the bacteria containing the IVT vector for mRNA were grown overnight from a glycerol stock in Gibco™ LB broth (1X) supplemented with 100 µg/mL Ampicillin with continuous shaking at 225 rpm in a 37°C incubator. The optical density (OD600) of the bacterial cells was 2.3 after 16 hours, as measured by a Thermo Scientific™ NanoDrop™ Spectrophotometer. The pDNA isolation was performed on a 96-well plate using the Applied Biosystems™ MagMAX™ Pro HT NoSpin Miniprep

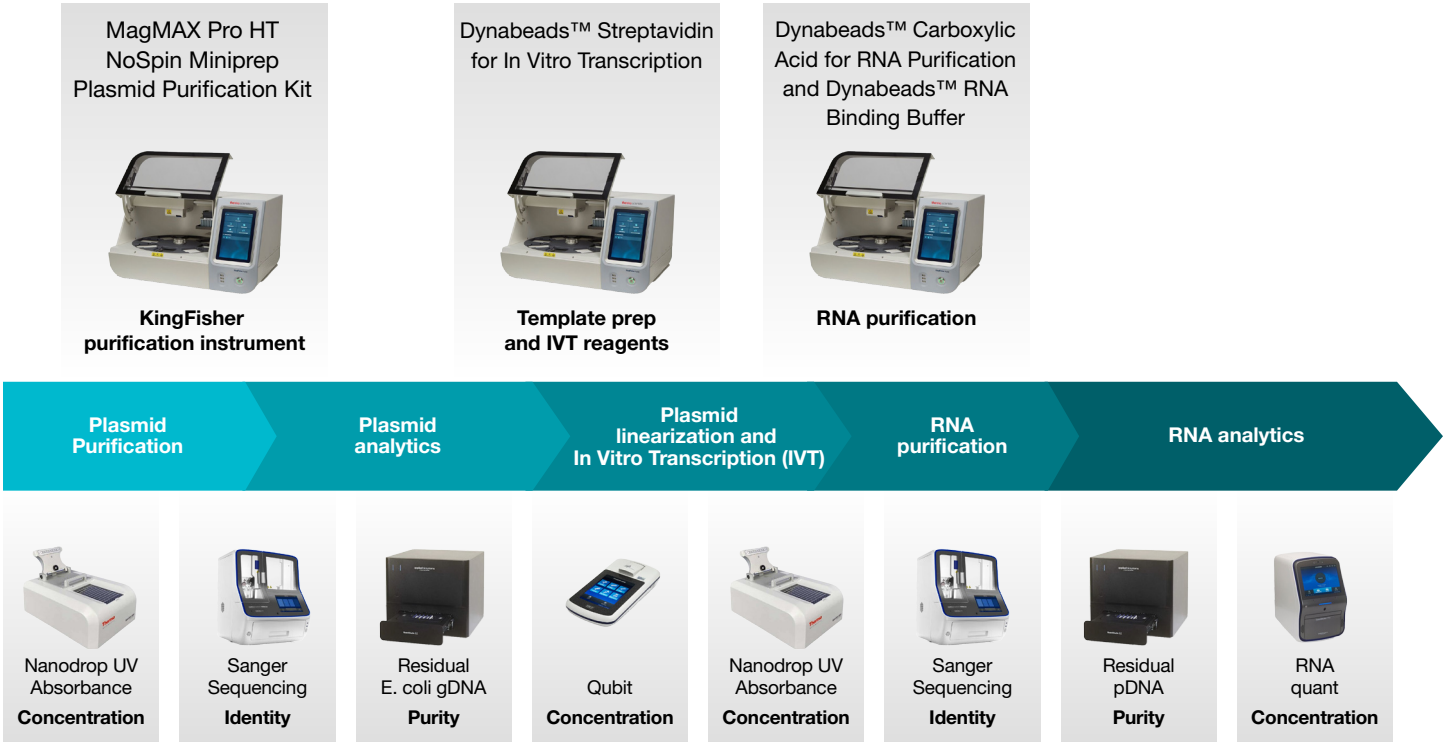


Figure 1: Overview of automated workflow for plasmid purification, *in vitro* transcription, mRNA purification and QC testing

Plasmid Purification Kit and the KingFisher Apex™ purification instrument. The workflow requires minimal user intervention and allows the processing of 96 samples in parallel in 32 minutes with additional 10 minutes hands-on time. (Figure 3).

Plasmids were linearized by digestion using LguI restriction enzyme and biotinylated by fill-in reaction of biotin-dUTP. 0.64 µg of linearized and biotinylated plasmid was then immobilized on 1 mg Invitrogen™ Dynabeads™ Streptavidin for In Vitro Transcription (Figure 4). Solid-phase IVT was performed using T7 RNA polymerase in 100 µL reactions for 2 hours using the Thermo Scientific™ TheraPure™ product portfolio of raw materials. After the IVT, the DNA-bead complex was removed from the reaction mix by applying a magnet.

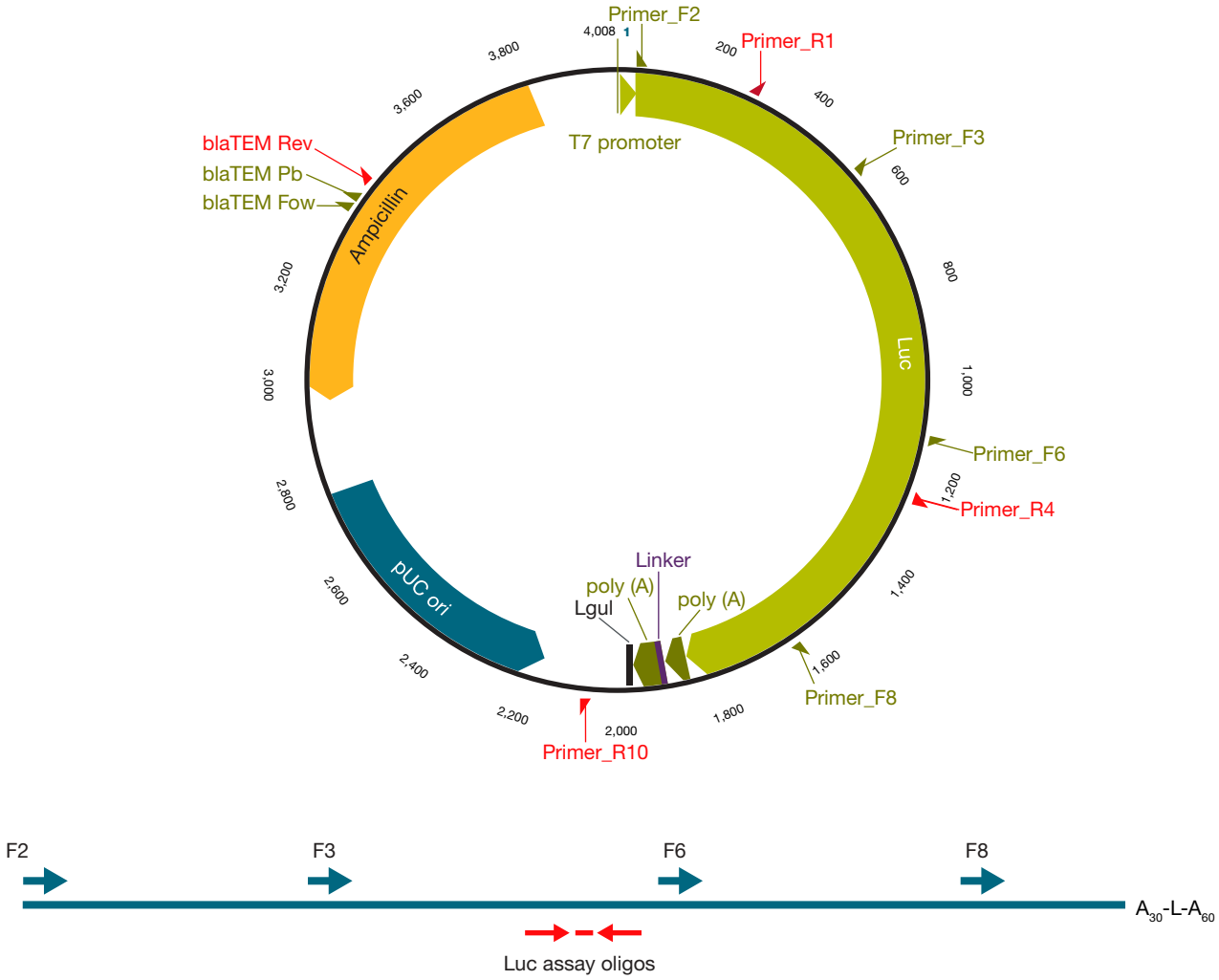


Figure 2: Map of pT7[mRNA]-{5utr-FLuc-3utr-pA}) Plasmid vector (A) and schematic view of RNA with Primer binding regions (B)

A: Location of T7 promoter (mint green), luciferase coding region (light green), segmented poly(A) tail (dark green) and in between linker (ochre), origin of replication (blue) as well as Ampicillin resistance marker (yellow) are indicated. Forward primers used for sequencing (F2, F3, F6, F8) are shown in dark green, while reverse sequencing primers are depicted in red (R1, R4, R10). The location of the dPCR assay oligos targeting the ampicillin resistance marker (blaTEM) for residual plasmid quantification are also shown. LguI restriction enzyme for plasmid linearization is located at the end of the poly(A) sequence. Plasmid figure derived from Geneious Prime software.

B: Location of sequencing primers (F2, F3, F6, F8) and the luciferase assay oligos for mRNA quantification are indicated. A_{30} and A_{60} refer to a stretch of 30 and 60 adenosines, respectively, in the segmented poly(A) tail, with the intervening linker (L).

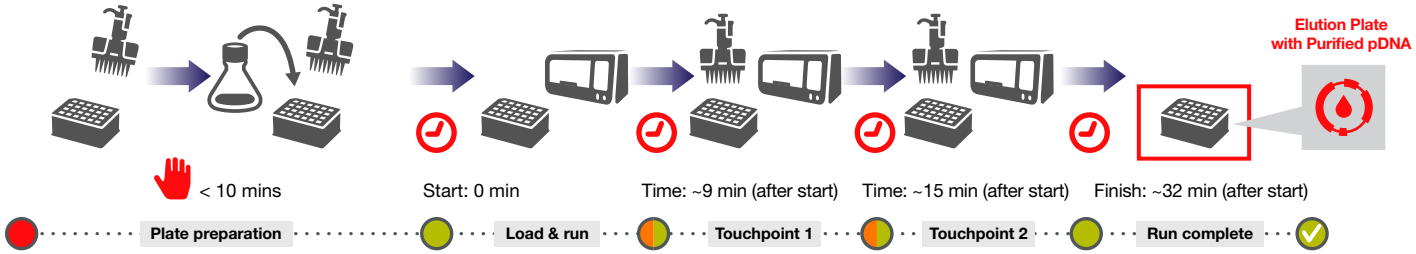


Figure 3: Benchtop extraction workflow for pDNA purification using a KingFisher system and the MagMAX Pro HT NoSpin Miniprep Plasmid Purification Kit

Schematic view of the workflow. In the plate preparation step, the operator prepares 5 plates with reagents (elution buffer, wash buffer 1 and 2, resuspension and lysis buffer) and 2 tip comb on two different 96-deep well plates. Prior to loading the plates on the instrument, BactBind beads are added to each well of the sample plate with 1ml of overnight bacterial cultures. After loading and initiating the run, the operator prepares the precipitation buffer and adds it to the 96-deep well lysate plate, and loads it back on the instrument in touchpoint 1. In touchpoint 2, the operator adds Endotoxin removal buffer and PlasmidPure beads simultaneously to the lysate plate, and returns the lysate plate onto the instrument. After the run, the operator retrieves the elution plate with the purified pDNA.

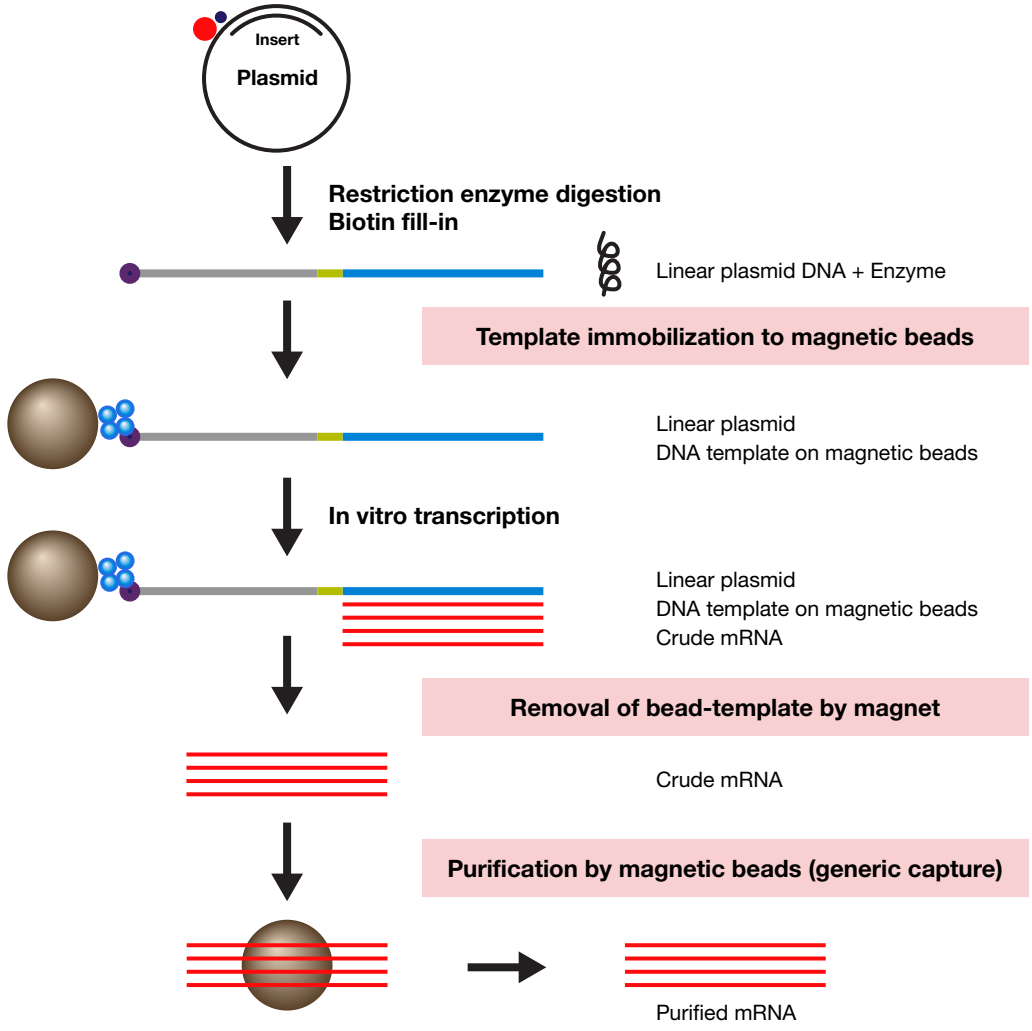


Figure 4: Solid-phase RNA production workflow on KingFisher instruments

Schematic view of the solid-phase IVT workflow on KingFisher. Linearized and biotinylated DNA template is immobilized to Dynabeads Streptavidin for IVT. The bead-DNA complex is used directly in IVT reaction. Following IVT, the template is removed from the synthesized mRNA. After magnetic removal of the DNA template, remaining IVT reagents are removed by generic capture of the mRNA onto Dynabeads Carboxylic Acid for RNA purification beads and Dynabeads RNA binding buffer.

Further, generic capture purification of RNA was performed using Invitrogen™ Dynabeads™ Carboxylic Acid for RNA purification magnetic beads together with Invitrogen™ Dynabeads™ RNA binding buffer (Figure 4). 100 µL of RNA was mixed with magnetic beads followed by adding 100 µL of RNA Binding Buffer. After washing, purified RNA was eluted in 100 µL of elution buffer, providing high quality mRNA. Template immobilization, solid-phase IVT and generic capture purification was performed on the Thermo Scientific™ KingFisher™ Apex instrument. For this protocol the RNA products were not capped.

Plasmid DNA and RNA concentration and purity by UV and Qubit

The concentration and purity of the purified DNA and mRNA was assessed by UV spectrometry on the nanodrop instrument according to standard methods. Concentration of crude RNA is determined using the Invitrogen™ Qubit™ RNA BR (Broad-Range) Assay Kit.

Sanger sequencing of plasmid DNA and mRNA

Sanger sequencing was performed using the Applied Biosystems™ BigDye™ Terminator v3.1 Cycle Sequencing Kit reagents on the luciferase coding region of the plasmid including the poly(A) tail. Supercoiled plasmids (5µL equal to 400-600ng) were directly added to the cycle sequencing reaction without any further preparation. Eight replicate plasmid preparations were cycled in separate reactions with four forward and three reverse primers in 1.5µL 5X Sequencing Buffer and 1µL Ready Reaction Mix. Four forward primers were spaced by 500-600 bases (F2, F3, F6, F8), while a reverse primer was added on each end and in the middle to cover a challenging region, see Figure 2 (R1, R4, R10).

A positive control was set up alongside with pGEM™-3Zf(+) plasmid and M13 forward and reverse primers (provided by the kit) to monitor the performance of the sequencing run. The reactions were subjected to BigDye XTerminator Purification Kit and then analyzed on the Applied Biosystems™ SeqStudio™ Flex Genetic Analyzer. Traces were analyzed with the InnoviGene™ Suite software v1.0 using the KB 1.4.2.6 base caller. First, quality metrics (trace score, CRL and QV20+) were evaluated against the acceptance limits listed in Table 1. The sequences were aligned to the plasmid sequence. For any mismatches or mixed bases found in the alignment of the sequences the corresponding traces were inspected. The % identity was calculated as quotient of the total number of divergent positions subtracted from the consensus sequence length over total consensus sequence length.

For RNA sequencing, cDNA was generated first with 500ng of RNA using the Invitrogen™ SuperScript™ IV VILO™ Master Mix. Eight RNA preps and a positive control RNA were *in vitro* transcribed. Negative control reactions without reverse transcriptase enzyme were set up alongside to monitor the presence of residual plasmid. After the cDNA synthesis, residual primers and nucleotides were hydrolyzed with ExoSAP-IT™. One microliter of the resulting cDNA was then subjected to cycle sequencing with each of the forward primers (see Figure 2). The reactions were processed afterwards as described for DNA sequencing and aligned to the 1960 bases of the plasmid sequence starting from the transcription start to the end of the poly(A) tail. The primers F6 and F8 produced shorter reads because of the vicinity to the end of the RNA. Therefore, the trace score, CRL and QV20+ acceptance limits listed in Table 1 were not applied to these.

Table 1. Preliminary acceptance limits for sample and plate validity

Test	Metric	Acceptance limit/ Range
Sanger sequencing	Trace score	>35
Sanger sequencing	CRL	>700
Sanger sequencing	QV20+	>700
dPCR pDNA quantification	Concentration of negative control	<0.3 copies/µL
dPCR pDNA quantification	Total valid events	>20,000
dPCR pDNA quantification	Concentration of internal control	+/-30% of calculated concentration
dPCR pDNA quantification	Concentration of positive control	+/-30% of calculated concentration
qPCR quantification	Negative control Cq	>35
qPCR quantification	PCR efficiency	90 - 110%.
qPCR quantification	R2 of standard curve	>0.99
qPCR quantification	Internal control Cq	28.8 - 31.8

Proposed preliminary acceptance limits are modified from (Hays et al., 2022; Hays et al., 2024) and derived from internal evaluationß

Residual *E. coli* gDNA and residual Plasmid detection by dPCR

Residual *E. coli* gDNA content was measured to determine the purity of the plasmid mini preparations. We employed the validated Applied Biosystems™ resDNASEQ™ dPCR *E. coli* DNA Kit on the Applied Biosystems™ QuantStudio™ Absolute QTM dPCR system. A DNA standard comprised of *E. coli* gDNA is used as positive control. The assay targets a single copy gene in the *E. coli* genome for quantification of genome copies. A DNA dilution buffer is provided for sample dilution ensuring stability of low-level DNA copies. An internal positive control (IPC) assay along with a synthetic template is used to monitor potential PCR inhibition and is detected in the second channel (VIC) of the instrument. Plasmid preps were diluted hundredfold and 3 µL of samples were tested in triplicates according to the kit instructions with the following modifications. Hold times for denaturation and annealing during cycling were increased to 10, 30 seconds, respectively.

The data from each run were evaluated against the acceptance criteria, as outlined in the manual to determine validity of plates and samples. The copies of *E. coli* genomic DNA measured by dPCR were converted into mass using the molecular weight of the *E. coli* strain TOP10 (GCF_019599065.1).

To detect the amount of residual plasmid DNA in the purified RNA, a FAM assay targeting the TEM β-lactamase gene (BlaTEM), a common Ampicillin resistance marker on plasmids, was used (see Figure 2). The BlaTEM assay was duplexed with the control Applied Biosystems™ Truemark™ Xeno™ assay (Xeno assay, ID: Ac00010014_a1, VIC-labelled) targeting an artificial DNA template (Applied Biosystems™ TrueMark™ Xeno Control, kanamycin resistance) of which 1000 copies were spiked into each reaction to monitor PCR performance. A positive control was set up by XhoI digestion of pGEM™-3Zf(+) plasmid (included with the BigDye Terminator v3.1 Cycle Sequencing Kit). Positive and negative (no template) controls were included on every MAP16 plate to monitor PCR performance and potential contamination. RNA eluates were diluted to 1ng/µL using the DNA dilution buffer and 3µL sample was tested in triplicates in the Absolute Q DNA Digital PCR Master Mix on the Absolute Q dPCR system with the same cycling protocol as described above. The data from each run were evaluated against the acceptance criteria listed in Table 1. The copies of plasmid DNA measured by dPCR were converted into mass using the molecular weight of the plasmid.

RNA quantification by qPCR

An assay targeting the center of the luciferase gene coding sequence was designed (luc assay, FAM labelled, see Figure 2) and duplexed with the above mentioned internal control assay (Xeno assay, ID: Ac00010014_a1, VIC-labelled) using the TaqPath™ DuraPlex™ 1-Step RT-qPCR Master Mix reagents. The Xeno target (TaqMan™ Universal RNA Spike In/Reverse

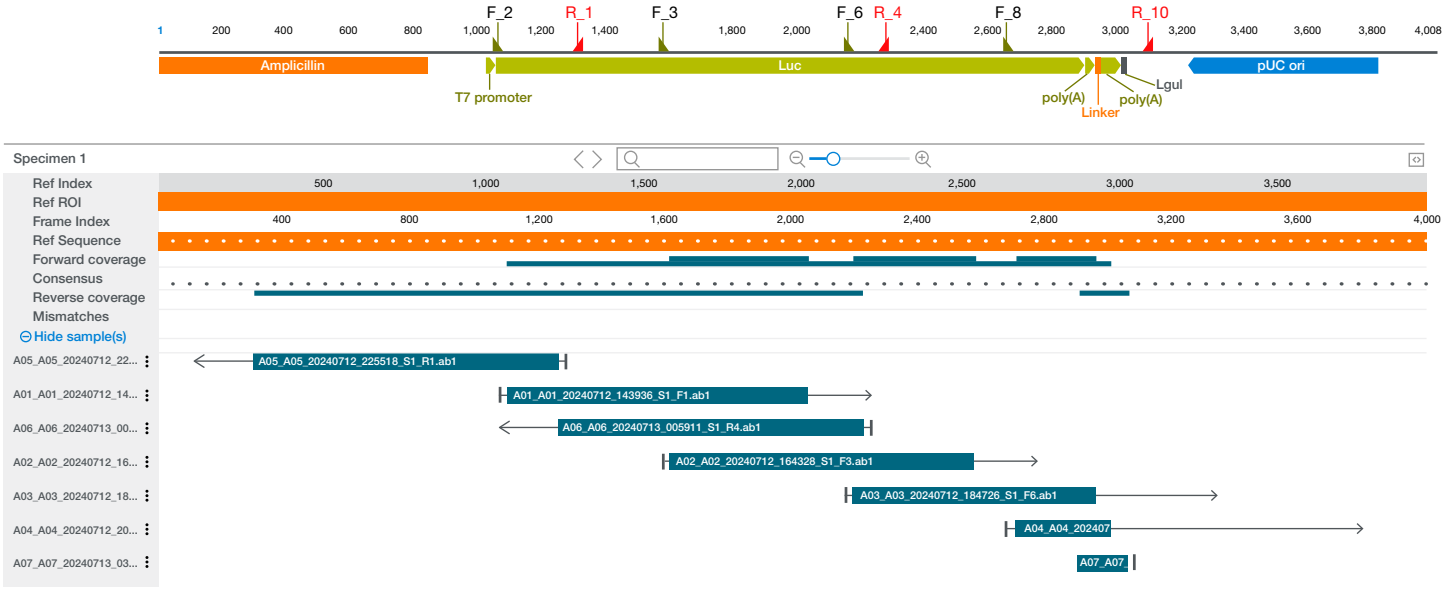


Figure 5: Alignment of plasmid DNA from Sanger Sequencing and plasmid reference

The top schematic view of the plasmid reference sequence (generated by Geneious Prime software) shows the location of the 7 sequencing primers corresponding to the start of the reads in the alignment below. The alignment, generated by InnoviGene™ Suite software, depicts the following tracks from the top to bottom: the plasmid reference sequence (orange bars), the coverage by reverse and forward primers (blue bars) and consensus sequence (dotted line), any mismatches, and the 7 reads (blue bars with arrows) of sample 1 aligned to reference sequence. The arrows indicate the direction of the reads and the thin bars on both sides represents the trimmed ends. Note that the track for mismatches is empty since there were none in this sample.

Transcription {Xeno} Control) was spiked in all reactions at 2000 copies as internal control. The primer and probe concentrations were optimized in the duplex reaction to reduce competition between the abundant luciferase and low copy Xeno target. The optimal concentrations for the luc assay were 300 nM primers and 250 nM probe, while the primer-limited concentration format (150 nM primers, 250 nM probe) worked best for the Xeno assay. A standard curve was set up using the TaqMan control dilution buffer and an mRNA transcript quantified by dPCR, with tenfold dilutions covering 7 orders of magnitude ranging from 10^9 to 10^2 copies per reaction. The purified RNA preparations were diluted to 1 ng/µL in the same buffer and tested in triplicates. The reactions were run with standard thermocycling conditions on the Applied Biosystems™ QuantStudio™ 7 Pro Real-Time PCR System. Data were analyzed using the Quantstudio Design & Analysis software. The input copies of the standards were directly entered into the software to obtain the regression analysis of the standard curve and quantified copies for the unknown samples. The data from each run were examined to ensure that they meet the specified preliminary acceptance criteria listed in Table 1.

Table 2: Sanger sequencing QC metrics for pDNA sample 1 and positive controls

Name	Status	Trace score	CRL	QV20+
A01_S1_F1.ab1	PASS	46.69	963	990
A02_S1_F3.ab1	PASS	46.89	989	987
A03_S1_F6.ab1	PASS	50.36	762	826
A04_S1_F8.ab1	PASS*	51.22	307	339
A05_S1_R1.ab1	PASS	42.97	971	973
A06_S1_R4.ab1	PASS	46.22	964	988
A07_S1_R10.ab1	PASS*	36.57	112	177
A08_pGEM_F.ab1	PASS	49.8	947	983
B08_pGEM_R.ab1	PASS	48.65	971	990

The QC metrics are defined as follows: Trace score: average of basecall quality values for bases in the clear range); QV20+: total number of bases in the entire trace that have a basecaller quality value ≥20, with a quality value of 20 having a probability of 1error in 100; CRL (contiguous read length): *CRL and CRV20+ exceptioned for short reads from F8 and R10 due to vicinity to polyA stretch. Cells highlighted yellow indicate suspect readout metrics, and cells highlighted red indicate a fail. The longest uninterrupted stretch of bases with quality values (QV) greater than or equal to a specified threshold.

Table 3: Summary sequence identity of plasmids

Specimen	Length	Forward	Reverse	Total mismatches	Identity*
Specimen 1	2752	4	3	0	100.00%
Specimen 2	2731	4	3	3	99.89%
Specimen 3	2731	4	3	0	100.00%
Specimen 4	2663	4	3	1	99.96%
Specimen 5	2748	4	3	0	100.00%
Specimen 6	2705	4	3	0	100.00%
Specimen 7	2706	4	2	0	100.00%
Specimen 8	2670	4	3	0	100.00%

*Calculated per correct bases over contig length

Table 4: List of mismatches in Sanger sequencing of plasmids

Specimen	Base change	Position	Type	Length	Location
Specimen 2	delC	375	DEL	1	Upstream of T7 promoter
	C > Y	880	SUB	1	Upstream of T7 promoter
	delT	3043	DEL	1	Downstream of poly(A) tail
Specimen 4	insT	483	INS	1	Upstream of T7 promoter

T7 promoter is located at 1034-1052, poly(A) at 2911-3010 of reference sequence.

Results

Automated plasmid DNA purification on the KingFisher system

The automated plasmid miniprep workflow was performed on the KingFisher Apex instrument (Figure 3). Notably, the plasmid purification on a 96-well plate was completed in 32 minutes (with two user touchpoints) and required less than 10 minutes of hands-on time. After the run, the DNA purity and yield were assessed by UV absorbance. The average plasmid DNA yield, A260/A280, and A260/A230 from a 96-well plate were 11 µg (SD + 2 µg), 2.22 (SD + 0.08), and 2.22 (SD + 0.29) respectively. The purity of the samples, evaluated through the A260/280 ratio, indicating protein contamination where a value >1.8 is considered pure and 260/230 ratio indicating carbohydrate or other contaminants where a value of >2 is considered pure was found to be acceptable. The plasmid quality was further tested using gel electrophoresis. The migration pattern indicates that the plasmids are supercoiled (data not shown).

Plasmid Sequence confirmation by Sanger sequencing

To ensure that the plasmids did not undergo any sequence alterations to the region of interest (ROI) while being propagated in *E. coli*, Sanger Sequencing was performed on eight of the plasmids. Four forward and three reverse primers were designed to cover the critical region starting from the T7 promoter to the segmented poly(A) tail (Figure 2). The sequencing reactions produced high quality reads. Table 2 shows exemplary QC data for the first sample with high trace scores (>35), high CRL and QV20+ values (>700), passing the preliminary acceptance criteria listed in Table 1 with the exception of reads from primers F8 and R10, with close vicinity to the poly(A) sequences. It is common that the sequence deteriorates after a long homopolymer stretch due to polymerase slippage.

The sequence reads were aligned to the reference plasmid sequence, as shown in Figure 5 for the first sample. Table 3 lists the summary of the sequencing run. All samples were confirmed to be >99.8% accurate. A total of four mismatches were found in 2 samples, as summarized in Table 4. Three of them were located >150 bases upstream of the T7 promoter and one of them 33 bases after the poly(A) tail. These mismatches are most likely sequencing errors occurring at the end of the sequencing traces and not a concern because they are not within the critical region between the T7 promoter and the end of the poly(A) tail.

Importantly, the entire sequence of the poly(A) tail was confirmed to be accurate in all eight samples. Figure 6 shows exemplary data for the first sample, with the traces showing the peaks for the segmented poly(A) tail, which is challenging to resolve due to the polymerase slippage. These data show that pDNA generated by the KingFisher MagMAX Pro HT NoSpin Miniprep Kit can be sequenced with a simple yet robust procedure to confirm the identity of sequences.

Residual *E. coli* gDNA content in Plasmid preparation

It is important that *E. coli* genomic DNA is removed during the plasmid purification process since the residual host DNA may decrease the performance in downstream applications. The Applied Biosystems™ resDNASEQ™ dPCR E. coli DNA Kit was employed to determine residual *E. coli* gDNA content in the first eight plasmid samples. The kit was designed for the Absolute Q dPCR system and contains all the necessary standards and controls to ensure an accurate, reliable measurement. Figure 7A and B show the fluorescence 1D scatter plots for the first three samples (A03, B03, C03) and controls. Positive fluorescence clusters for the *E. coli* assay and the internal positive control (IPC) assay were well separated from negative clusters (Figure 7A and B).

Figure 6: Sequencing of poly(A) region

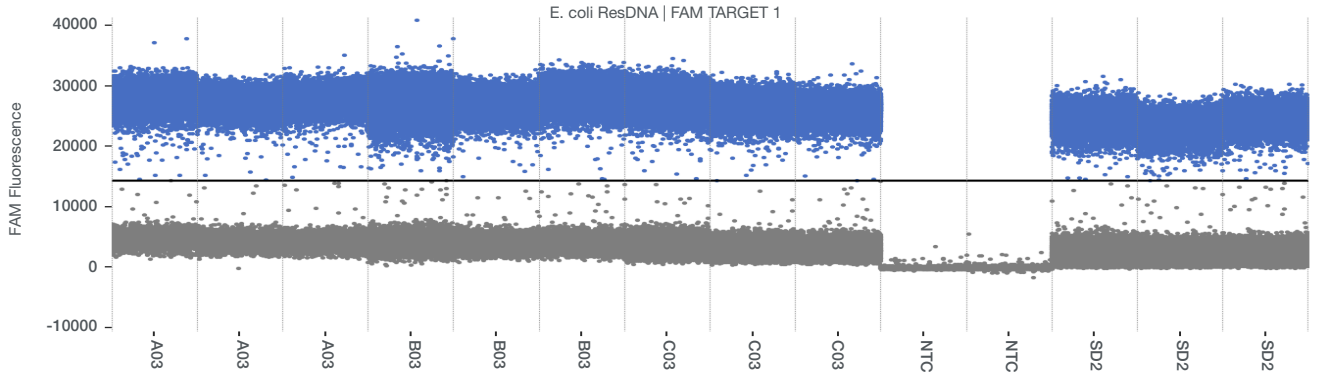
The InnoviGene™ Suite software was used to visualize the sequence alignment for the poly(A) region. Shown are traces for forward primers F6, F8 and reverse primer R10. The traces of forward primers show lower intensity and fade out due to the challenging poly(A) stretch, but read the 10 base linker correctly. The reverse sequence trace derived from primer R10 reads through the entire poly(A) sequence.

The runs passed the acceptance criteria for total events and controls listed in the instruction manual. The CV of the replicates for the E. coli assay concentration were less than 2.5% for all samples (Figure 7C), indicating high precision of the measurement. The *E. coli* gDNA content was determined to be on average at 5.1% ranging from 3.4-6.7% of the total DNA measured by nanodrop (Figure 7C). These data suggest that the plasmid DNA is suitable for downstream applications, like *in vitro* transcription.

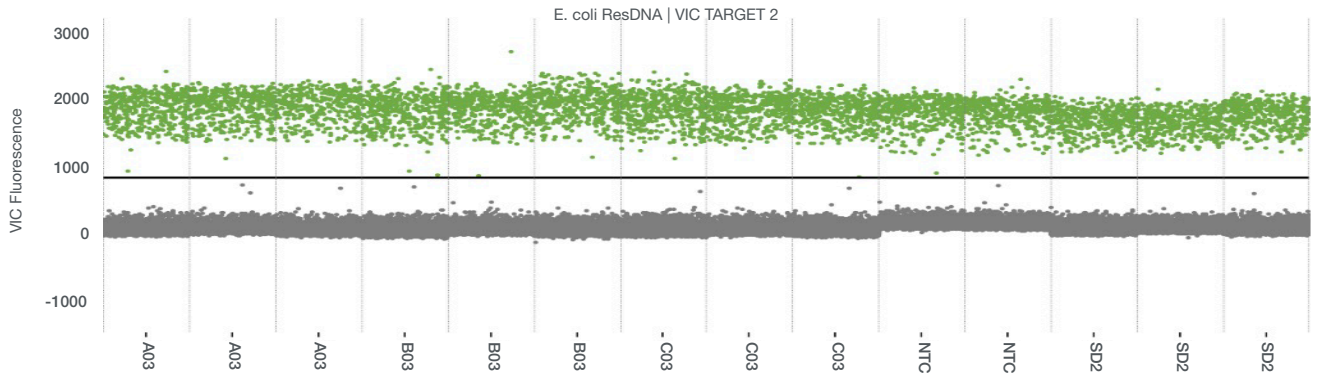
In vitro transcription and purification of RNA on KingFisher Apex

Solid-phase *in vitro* transcription of eight plasmid samples and generic capture purification of the mRNA was performed on the KingFisher Apex instrument (Figure 4). The automated workflow generated a high quality of poly(A) tailed mRNA, while cumbersome steps from conventional protocols are eliminated, i.e. removal of supercoiled plasmid and DNase treatment

A



B



C

	A260	dPCR				
	Stock	dPCR reaction			Stock	*E. coli* DNA content
	total DNA	*E. coli* gDNA		IPC	*E. coli* gDNA	
Sample	ng/µL	copies/µL	CV [%]	copies/µL	ng/µL	%
A03	128	5082.53	1.17	48.87	7.43	5.8%
B03	123.7	3945.22	0.17	51.19	5.77	4.7%
C03	123.5	2877.84	0.70	53.89	4.21	3.4%
D03	119	3730.31	2.12	49.24	5.45	4.6%
E03	120.7	4511.72	0.95	50.77	6.59	5.5%
F03	125.8	4225.12	1.38	50.13	6.17	4.9%
G03	116.5	3953.8	0.66	52.37	5.78	5.0%
H03	139.3	6370.65	0.68	51.89	9.31	6.7%
					Average	5.1%
					Min	3.4%
					Max	6.7%

Figure 7: *E. coli* residual gDNA from 8 representative samples from Plasmid DNA purification

A depicts a 1D fluorescent scatter plot for the FAM-labelled E. coli assay (detected in Channel 1; positives marked in blue) and B for the VIC-labelled IPC assay (detected in Channel 2; positives marked in green). The plots show representative data for the first three samples A03, B03, C03 in triplicates along with the non-template control (NTC) and positive control (SD2). Note that positive and negative fluorescence clusters are well separated. C shows the dPCR quantification summary of all eight samples, including the IPC (internal positive control) along with the nanodrop reading of the plasmid concentration and the calculated E. coli DNA in the plasmid stocks (concentration and relative % by weight).

following the IVT step. Table 5 lists the concentration of the crude and purified RNA. The recovery of purified RNA was determined to be an average of 104%. The purified RNA had excellent purity according to A260/A280 and A260/A230 ratios. Furthermore, the purified RNA had >98% protein removal measured by modified micro BCA assay and high RNA integrity analyzed using the Agilent TapeStation (data not shown).

In summary, the automated, solid-phase IVT workflow enables high throughput screening, including template immobilization, IVT and mRNA purification of 96 samples in less than 5 hours. The workflow is simple, flexible and highly scalable, and template reuse enables low manufacturing footprint. Importantly, the same technology can be employed from research and development to vaccine manufacturing.

mRNA sequence confirmation by Sanger sequencing

The sequence of the *in vitro* transcribed RNA was confirmed by sanger sequencing of the cDNA using the same forward primers as for the plasmid sequencing (Figure 2). All sequencing reactions met the predetermined quality control criteria, with trace score values higher than 35 and CRL and CV20+ values exceeding 700, except the reads from primer F8, which is reading across the poly(A) tail and thus produces shorter CRL. Table 6 provides exemplary QC data for the first sample. No traces were found in the no-RT control reactions (data not shown), indicating that sequencing reactions were based on cDNA and did not pick up any residual pDNA in the samples.

Table 5: Crude and purified RNA concentration

	Qubit RNA BR Assay kit		% Recovery purified RNA *	NanoDrop measurements for purified RNA		
Sample	Crude RNA purification mix (RNA input – µg/µL)	Purified RNA (µg/µL)	(Purified RNA /Crude RNA purification mix) x 100	Yield (µg/µL) (A260nm)	A260/A280	A260/A230
S1	0.87	0.92	106	1.21	2.17	2.49
S2	0.83	0.84	101	1.16	2.18	2.51
S3	0.89	0.8	90	1.07	2.17	2.49
S4	0.73	0.77	105	1.04	2.16	2.49
S5	0.83	0.86	103	1.13	2.17	2.50
S6	0.75	0.77	103	1.05	2.15	2.48
S7	0.87	0.93	107	1.24	2.17	2.50
S8	0.76	0.88	116	1.17	2.17	2.49

* Purified RNA percent recovery is calculated from the Qubit RNA BR Assay Kit readings

Table 6: QC metrics for RNA (cDNA) sequencing

Name	Status	Trace score	CRL	QV20+
A01_S1_F2.ab1	PASS	45.98	976	975
A02_S1_F3.ab1	PASS	49.4	943	985
A03_S1_F6.ab1	PASS	51.41	743	784
A02_S1_F8.ab1	PASS*	44.09	241	337

CRL, continuous read length, QV20+

The sequences were aligned to a reference sequence as shown in Figure 8A for the first sample, resulting in a consensus sequence covering almost the entire reference sequence (Figure 8B). Table 7 shows a summary of the alignments for all samples. The verified high-quality sequences started around 35-40 bases downstream of the forward primer (located at position 1-23 in the reference sequence) and covered the majority if not all of the poly(A) tail, confirming its presence. Eight out of eight samples matched the reference sequence with identities exceeding 99.9%. The mismatch in two samples was located right at the start of the first poly(A) stretch (Table 8), most likely due sequencing errors caused by polymerase slippage. These data show that the sequence of the RNA can be verified using a direct sequencing approach with cDNA, which is a faster workflow than methods relying on PCR amplification of the cDNA. Avoiding the additional PCR step not only shortens the protocol by approximately 1.5 hours, but it also reduces the risk of introducing errors due to misincorporation of bases in PCR.

Table 7: Sanger sequencing summary for RNA (cDNA)

Sample	contig length [bp]	Ref start	Ref end	Forward	Total mismatches	Identity*
1	1898	63	1960	4	0	100.00%
2	1902	59	1960	4	1	99.95%
3	1902	59	1960	4	0	100.00%
4	1898	63	1960	4	0	100.00%
5	1905	56	1960	4	0	100.00%
6	1888	59	1946	4	0	100.00%
7	1877	59	1934	4	1	99.95%
8	1902	59	1960	4	0	100.00%

*Calculated per correct bases over contig length

Poly(A): 1861-1960

Table 8: Mismatches in RNA

Specimen	Base change	Position	Type	Length	Location
Specimen2	A > C	1861	SUB	1	First base in Poly(A)
Specimen7	insC	1860	INS	1	Last base before Poly(A)

Poly(A): 1861-1960

Residual plasmid DNA detection in mRNA by dPCR

Removal of the template DNA and other impurities from mRNA is essential to achieve high mRNA translation levels and prevent the activation of undesirable immune responses (Zhang et al., 2023). Levels of residual plasmid DNA in the mRNA were determined by dPCR on the Absolute Q dPCR system. An assay targeting the ampicillin resistance marker on the plasmid was duplexed with an artificial internal control assay (Xeno assay) to monitor for PCR inhibition. The duplex assay showed well separated fluorescence clusters (Figure 9A and B). Total valid events as well as positive, negative and internal controls passed the acceptance criteria outlined in Table 1.

As shown in Figure 9C, the residual plasmid DNA in all RNA preps was <10ng/mg of RNA, with an average of 7.79 and a range from 6.45-9.23 ng/mg RNA.

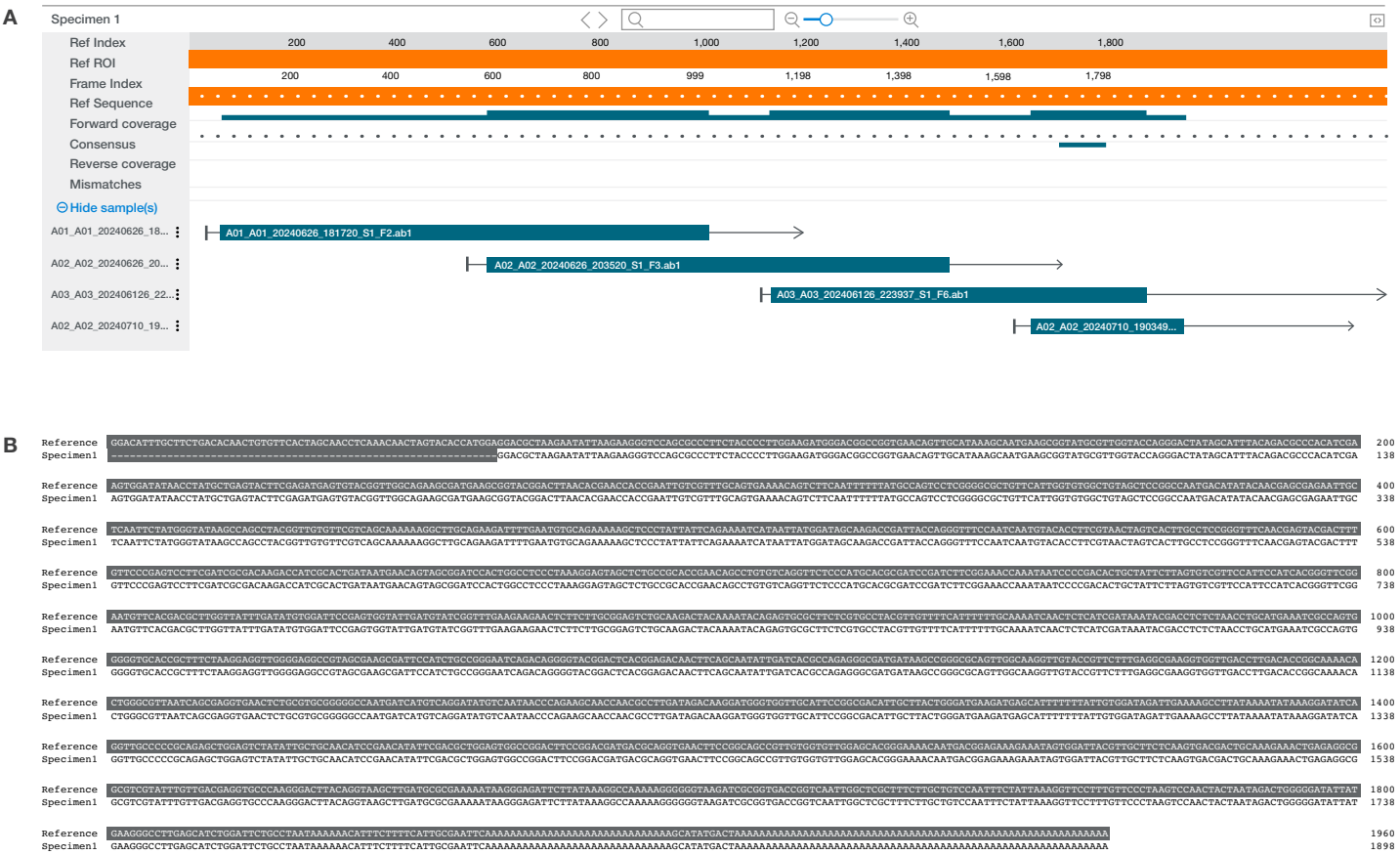
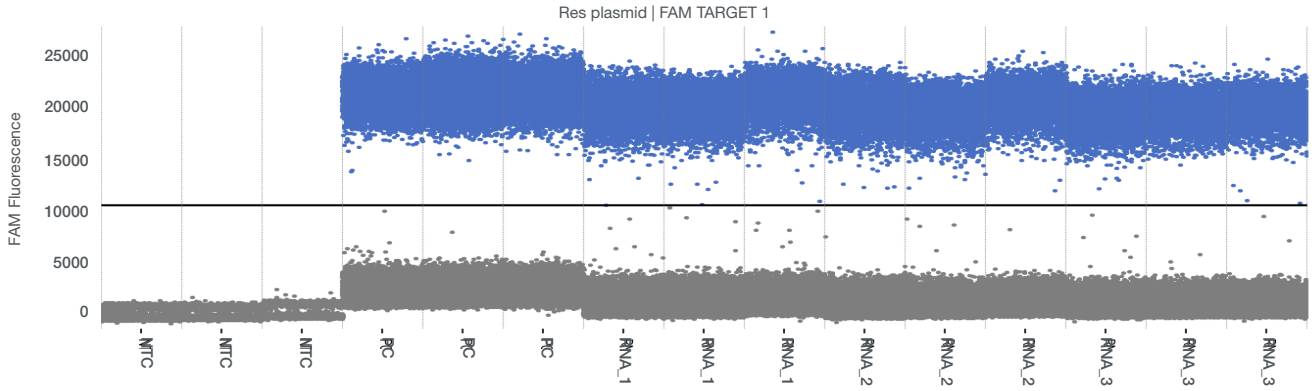


Figure 8: Sanger sequencing alignment of RNA (cDNA) for first sample

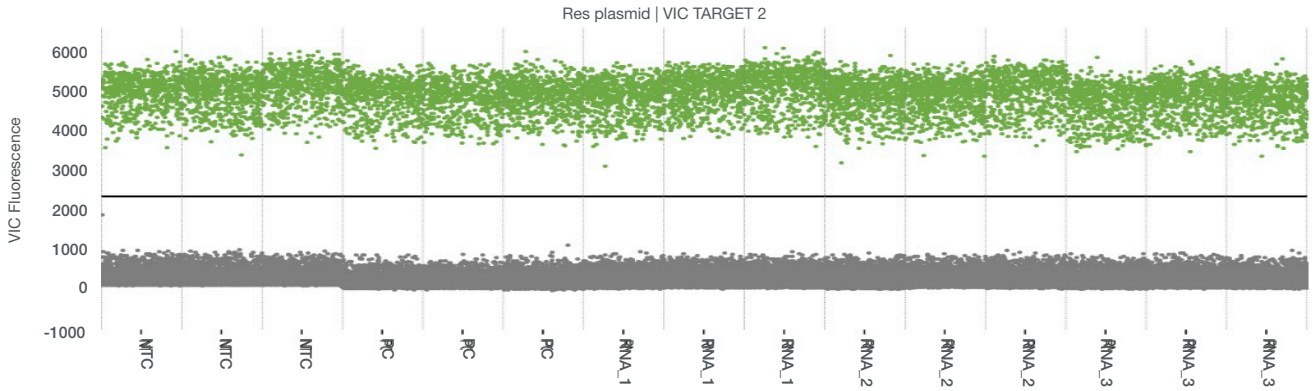
A: Sequence alignment of the first sample composed of the reference mRNA sequence (orange) and forward primers F2, F3, F6, F8 generated by the InnoviGene™ Suite software. Shown are the coverage (blue), and the 4 reads (blue bars with arrows), of sample 1 aligned to reference sequence. The arrows indicate the direction of the reads and the thin bars on both sides represents the trimmed ends. Note that the track for mismatches is empty since there were none in this sample.

B: Sequence alignment of reference and consensus sequence for sample 1. Note that the entire sequence starting from position 63 until the very end of the poly(A) tail (position 1960) was confirmed.

A



B



C

	A260-RNA	dPCR pDNA		DNA content
	Stock	dPCR reaction		
	Total RNA	Plasmid		
Sample	ng/µL	copies/µL	CV [%]	ng DNA/mg RNA
RNA_1	1.21	711.68	0.19	9.24
RNA_2	1.16	585.46	1.67	7.60
RNA_3	1.07	497.18	1.37	6.45
RNA_4	1.04	597.35	1.33	7.75
RNA_5	1.13	518.9	1.28	6.73
RNA_6	1.05	612.96	1.41	7.96
RNA_7	1.24	631.83	0.86	8.20
RNA_8	1.17	648.41	1.10	8.42
			Average	7.79
			Min	6.45
			Max	9.24

Figure 9: Residual plasmid DNA from 8-representative samples from purified RNA

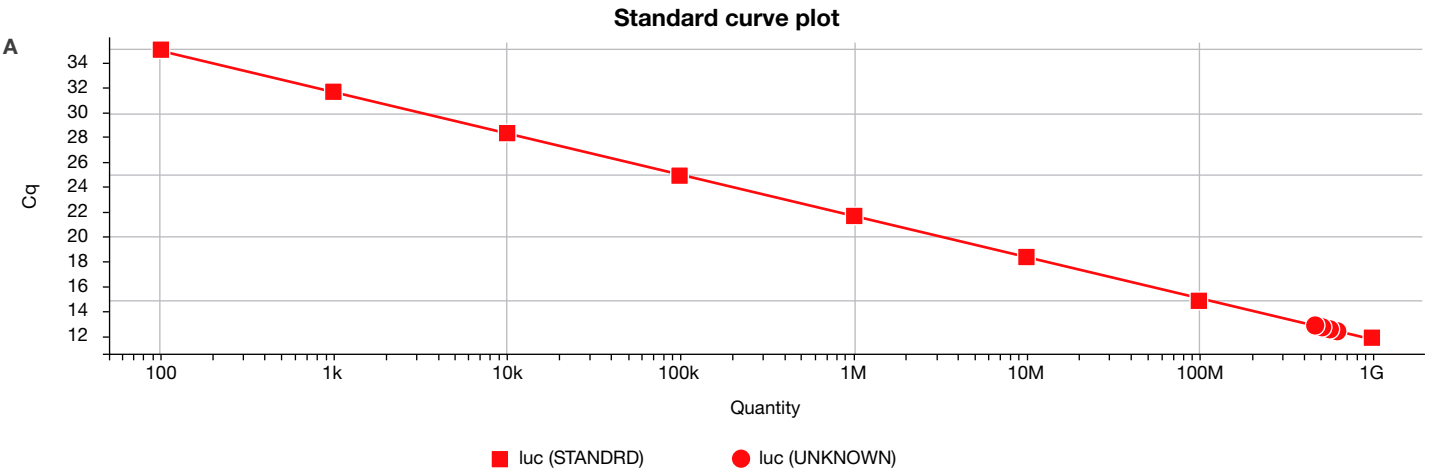
A and B: The FAM-labelled BlaTEM assay (detected in Channel 1; positives marked in blue) and the VIC-labelled xeno assay (detected in Channel 2; positives marked in green) show ample cluster separation in the 1D fluorescent scatter plots. Shown are representative data for the non-template control (NTC), positive control (PC) and the first three RNA samples (RNA_1-3) in triplicates. C shows the dPCR quantification summary of all eight RNA samples along with the nanodrop reading of the RNA concentration and the calculated plasmid DNA content in the RNA stocks.

mRNA quantification by qPCR

Finally, the mRNA was quantified by qPCR. A duplex assay comprised of a luciferase gene assay and an internal control assay was set up using the Applied Biosystems™ TaqPath™ DuraPlex™ 1-Step RT-qPCR Master Mix. This mastermix was developed for multiplexing of up to six targets and is automation-friendly due to low viscosity and long benchtop-stability for fully assembled reactions.

Eight purified RNA samples were tested with the luciferase/ Xeno duplex qPCR assay. After the run, the data were evaluated against preliminary acceptance criteria (Table 1). All listed criteria were met. The standard curve for the luciferase assay had excellent linearity as shown in Figure 10, with 100.12% PCR efficiency, and an R^2 value of 1, well within the preliminary acceptance range (90-110% efficiency, R^2>0.99). The internal control assay showed very similar Cq across all samples, indicating good PCR performance and absence of inhibition. The purified RNA (1ng/µL dilution) had a concentration of $4.90x10^8$ to $5.24x10^8$ copies/µL, see Figure 10, which is about 55% of the theoretical value based on molecular weight and RNA concentration per nanodrop reading (Table 5).

A



B

		Luciferase assay (luc) in FAM				Xeno assay in VIC	
Sample	Replicates	Average Cq	Stdev Cq	Quantity mean [copies]	Quantity SD	Average Cq	Stdev Cq
NTC	9	37.95	0.96	N/A	N/A	30.24	0.10
RNA_1	3	12.68	0.02	$5.11x10^8$	$6.38x10^6$	30.12	0.01
RNA_2	3	12.68	0.03	$5.10x10^8$	$1.09x10^7$	30.13	0.06
RNA_3	3	12.66	0.02	$5.16x10^8$	$8.52x10^6$	30.20	0.10
RNA_4	3	12.72	0.03	$4.97x10^8$	$8.68x10^6$	30.20	0.01
RNA_5	3	12.68	0.02	$5.11x10^8$	$5.85x10^6$	30.26	0.06
RNA_6	3	12.74	0.00	$4.90x10^8$	$1.37x10^6$	30.16	0.01
RNA_7	3	12.64	0.02	$5.24x10^8$	$6.90x10^6$	30.23	0.02
RNA_8	3	12.73	0.02	$4.94x10^8$	$7.77x10^6$	30.17	0.06

Figure 10: Quantification of purified RNA by qPCR

A A standard curve for the luciferase qPCR assay was generated using the Design and Analysis software. The starting quantities of the RNA standard, an 8-point dilution from 10^2 to 10^9 copies/reaction were directly entered into the software. The standard curve analysis module generated the plot of Cq vs Quantity and the standard curve with the following values: Slope: -3.316, intercept (Y-Inter): 41.581. This corresponds to a PCR efficiency of 100.121%. The R2 value was determined to be 1. The graph shows the standard curve generated from the data points of the standard (squares) and the unknown samples (circles).

B The table summarizes the Cq values for the luciferase and the xeno internal control assay as well as the quantity of luciferase RNA measured in the samples based on the standard curve.

Summary

The discovery and development of mRNA interventions is an exciting area of research. However, *in vitro* transcription and purification of mRNA can be a cumbersome workflow. We demonstrate a complete automated workflow with minimal hands-on time for plasmid sample preparation and template preparation, *in vitro* transcription and mRNA purification. The demonstrated workflow using magnetic beads to immobilize the DNA template, perform *in vitro* transcription and mRNA purification is easily automated for high throughput and high reproducibility. The bead-DNA complex could be stored for repeated *in vitro* transcription reactions.

We also describe analytical methods to assess the quality of purified pDNA and RNA. Specifically, we describe Sanger Sequencing protocols for supercoiled plasmid and cDNA of the RNA to confirm their identity. Additionally, we utilize dPCR assays to determine the purity by measuring residual host or plasmid DNA and quantify the concentration of RNA by qPCR.

qPCR and dPCR can be used interchangeably for these applications. However, dPCR may be more suitable for detection of low abundance targets due to the higher precision for low events. On the other hand, qPCR is the better choice for quantification if a wider dynamic range is needed (Hays et al., 2022).

The qPCR and dPCR workflows can be integrated in an automated environment if higher throughput is needed. For qPCR, the Thermo Scientific™ Orbitor™ RS2 Microplate Mover can be employed for plate loading onto the QuantStudio 7 Pro. Similarly, for dPCR, the Applied Biosystems QuantStudio Absolute Q AutoRun Digital PCR (dPCR) Suite provides flexible instrument configurations for one or two dPCR instruments and intelligent automation that supports up to 72 hours of hands-free operation.

The sequencing and PCR methods described herein can be easily customized to detect other targets using ThermoFisher oligo and assay design tools such as Invitrogen™ OligoPerfect™ and Custom TaqMan® Assay Design Tool. To facilitate the process for researchers, we provide convenient kits and reagents, as well as instrumentation and software that enable reproducible results for analytical testing at the early stages of the discovery process. Furthermore, these methods can be modified and validated for use in later development stages for in-process, lot release and QC testing of mRNA drug substances and products.

In conclusion, we provide a simple yet robust workflow for plasmid purification, RNA production and quality control of the purified nucleic acids. These scalable solutions provide high quality plasmids and RNA with minimal hands-on time, reducing the potential risk for errors. By adopting these methods, the pipeline for mRNA therapy production and QC can be streamlined and made available faster.

Ordering information

Description	Catalog number
Plasmid isolation, IVT and RNA purification	
MagMAX Pro HT NoSpin Plasmid Purification Kit	A58309
Dynabeads™ Streptavidin for In Vitro Transcription	65005D, 65006D
Dynabeads™ Carboxylic Acid for RNA Purification	65021D, 65020D
Dynabeads™ RNA Binding Buffer	65042D
MEGAscript™ T7 Transcription Kit	AM1333
Quality control instruments	
QuantStudio Absolute Q Digital PCR System	A52864
QuantStudio™ 7 Pro Real-Time PCR System, 96-well, 0.1 mL	A43163
SeqStudio 24 Flex Genetic Analyze	A53630
SeqStudio 8 Flex Genetic Analyze	A53627
SeqStudio Genetic Analyzer System	A35644
Software	
InnoviGene™ Suite v1.0	
Applied Biosystems™ QuantStudio™ Absolute Q Digital PCR Software Release Version: 6.3.3	
Design & Analysis Software Release Version:2.8.0	
Quality control reagents	
SuperScript™ IV VILO™ Master Mix	1175605
BigDye Terminator v3.1 Cycle Sequencing Kit	4337455
BigDye XTerminator Purification Kit	4376486
Applied Biosystems™ ExoSAP-IT™ PCR Product Cleanup Reagent	78201.1.ML
QuantStudio Absolute Q MAP16 Plate Kit	A52865
resDNASEQ™ dPCR E. coli DNA Kit	A57017
Absolute Q DNA Digital PCR Master Mix (5X)	A52490
Applied Biosystems™ TaqMan™ Gene Expression Assay (FAM), ID: Ba04646128_s1	4331182
Applied Biosystems™ TaqMan™ Gene Expression Assay, VIC, ID: Ac00010014_a1	4448489
Applied Biosystems™ TrueMark™ Xeno Control, kanamycin resistance	A50384
Applied Biosystems™ DNA Dilution Buffer	4405587C
Applied Biosystems™ TaqMan™ Gene Expression Assay, VIC primer-limited, ID: Ac00010014_a1	4448484
Applied Biosystems™ TaqMan™ Universal RNA Spike In/Reverse Transcription (Xeno) Control	A39179
TaqMan control dilution buffer	A49889

For Research Use Only. Not for use in diagnostic procedure

Description	Catalog number
Plasmid isolation, IVT and RNA purification	
KingFisher Flex Purification System	24074431
KingFisher Apex Purification System	5400930
Quality control reagents	
Applied Biosystems™ TaqPath™ DuraPlex™ 1-Step RT-qPCR Master Mix	A58666

For Laboratory Use

Hays, A., Islam, R., Matys, K., & Williams, D. (2022). Best Practices in qPCR and dPCR Validation in Regulated Bioanalytical Laboratories. *AAPS J, 24*(2), 36. https://doi.org/10.1208/s12248-022-00686-1

Hays, A., Wissel, M., Colletti, K., Soon, R., Azadeh, M., Smith, J., Doddareddy, R., Chalfant, M., Adamowicz, W., Ramaswamy, S. S., Dholakiya, S. L., Guelman, S., Gullick, B., Durham, J., Rennier, K., Nagilla, P., Muruganandham, A., Diaz, M., Tierney, C., . . . Johansson, O. (2024). Recommendations for Method Development and Validation of qPCR and dPCR Assays in Support of Cell and Gene Therapy Drug Development. *AAPS J, 26*(1), 24. https://doi.org/10.1208/s12248-023-00880-9

Metkar, M., Pepin, C. S., & Moore, M. J. (2024). Tailor made: the art of therapeutic mRNA design. *Nat Rev Drug Discov, 23*(1), 67-83. https://doi.org/10.1038/s41573-023-00827-x

Pardi, N., Hogan, M. J., Porter, F. W., & Weissman, D. (2018). mRNA vaccines – a new era in vaccinology. *Nat Rev Drug Discov, 17*(4), 261-279. https://doi.org/10.1038/nrd.2017.243

Parhiz, H., Atochina-Vasserman, E. N., & Weissman, D. (2024). mRNA-based therapeutics: looking beyond COVID-19 vaccines. *Lancet, 403*(10432), 1192-1204. https://doi.org/10.1016/S0140-6736(23)02444-3

Qin, S., Tang, X., Chen, Y., Chen, K., Fan, N., Xiao, W., Zheng, Q., Li, G., Teng, Y., Wu, M., & Song, X. (2022). mRNA-based therapeutics: powerful and versatile tools to combat diseases. *Signal Transduct Target Ther, 7*(1), 166. https://doi.org/10.1038/s41392-022-01007-w

Rosa, S. S., Nunes, D., Antunes, L., Prazeres, D. M. F., Marques, M. P. C., & Azevedo, A. M. (2022). Maximizing mRNA vaccine production with Bayesian optimization. *Biotechnol Bioeng, 119*(11), 3127-3139. https://doi.org/10.1002/bit.28216

To, K. K. W., & Cho, W. C. S. (2021). An overview of rational design of mRNA-based therapeutics and vaccines. *Expert Opin Drug Discov, 16*(11), 1307-1317. https://doi.org/10.1080/17460441.2021.1935859

Webb, C., Ip, S., Bathula, N. V., Popova, P., Soriano, S. K. V., Ly, H. H., Eryilmaz, B., Nguyen Huu, V. A., Broadhead, R., Rabel, M., Villamagna, I., Abraham, S., Raeesi, V., Thomas, A., Clarke, S., Ramsay, E. C., Perrie, Y., & Blakney, A. K. (2022). Current Status and Future Perspectives on MRNA Drug Manufacturing. *Mol Pharm, 19*(4), 1047-1058. https://doi.org/10.1021/acs.molpharmaceut.2c00010

Zhang, J., Liu, Y., Li, C., Xiao, Q., Zhang, D., Chen, Y., Rosenecker, J., Ding, X., & Guan, S. (2023). Recent Advances and Innovations in the Preparation and Purification of *In Vitro*-Transcribed-mRNA-Based Molecules. *Pharmaceutics, 15*(9). https://doi.org/10.3390/pharmaceutics15092182