Establishment of potential reference measurement procedure and reference materials for *EML*4-*ALK* fusion variants measurement

Yi Yang[1,2], Yu Zhang[1], Shujun Zhou[2], Xia Wang[1], Chunyan Niu[1], Yongzhuo Zhang[1], Lianhua Dong[1,*]

1. Center for Advanced Measurement Science, National Institute of Metrology, Beijing, 100029, China

2. Shenzhen Institute for Technology Innovation, National Institute of Metrology, Shenzhen 518107, China

[*]Corresponding author. Tel: +86 010 64524959; fax: +86 010 64524959

E-mail address: donglh@nim.ac.cn

In vitro transcription RNA preparation and validation

The pBluescrip II SK (+) plasmid containing *EML4-ALK* (containing V1 and V3) and partial *ALK* (NM_004304.5, domain at the junction of exon 20 and 21) was synthesized from BGI (Beijing, China) to generate in vitro transcribed RNA molecules. The plasmid (4402 bp) was linearized by Xba I (Takara, Code No. 1093S) according to the manufacturer's instructions. The linearized plasmid was purified and checked using a commercial DNA Clean-up Kit (CW2301M, CWBIO, Jiangsu, China) with 50 μL of elution buffer and 2% agar gel electrophoresis (Supplemental Figure 1a). The results showed that the band was single and of consistent length, allowing for further in vitro transcription. Subsequently, only the purified plasmid was added to the in vitro transcription mixture following the MEGAscript™T7 transcription kit instructions (Thermo Fisher Science, USA). The transcribed RNA was then purified with the MEGAclear™ kit (Thermo Fisher Science, USA). The quantity and quality of in vitro transcribed RNA were assessed using the Nanodrop 2000 (Thermo Fisher Science, USA) and Bioanalyzer (Agilent 2100, USA) with RNA 6000 Nano Reagents Part I kit (Agilent, USA), respectively. The result of Bioanalyzer showed only one major peak existed in this transcript (Supplemental Figure 1b). And then the undigested in vitro transcribed RNA was analyzed to check the nucleotide impurity by mass spectrometry and the result showed that no nucleotide impurity existed (Supplemental Figure 1c). Eventually, the in vitro transcribed RNA was verified the bases sequence by BGI Tech Solutions (Beijing Liu He) Co., Limited. There was no mismatch observed in the in vitro transcribed RNA sequencing results (Supplementary Figure 2).

Calculation of NMPs and RNA concentration

The final mass fraction of each NMP (W_x) in the digested in vitro transcribed RNA was calculated by plotting the standard curve (Supplemental Figures 5). The final mass fraction of the RNA concentration (W_{RNA}) is calculated according to equation (1), where M_{RNA} is the molecular mass of the RNA molecule (519166 was used), M_{NMP} is the molecular mass of the selected NMP, and N is the number of the selected NMP in the RNA sample (404 for AMP, 314 for UMP, 425 for GMP, 376 for CMP). The copy number of RNA (n, copies/μg) was converted the mass fraction of RNA (W_{RNA}, μg g^{-1}) by equation (2), where N_A is Avogadro's constant, D is the density of the sample. A density of 1.00 g/mL was taken into account in the calculation.

$$W_{RNA} = \frac{W_x \times M_{RNA}}{M_{NMP} \times N} \qquad (1)$$

$$n = \frac{W_{RNA} \times D \times N_A}{M_{RNA} \times 10^9} \quad (2)$$

False-positive rate (FPR) assessment

The FPR of dPCR platform was calculated from the non-targeted template of a highly concentrated pure sample only. Five commercial dPCR platforms were selected divided into two categories: water-in-oil droplet dPCR platform (QX200, Microdrop-100, DQ24) and microfluidic chip dPCR platform (QIAcuity One, BioDigital Yan). The QX200 and DQ24 of specific details are listed in Supplemental Table 2. 2 μL template was added to the 20 μL reaction system (one-step) of the Microdrop-100 platform, the thermal cycling consisted of a 30 min reverse transcription at 55℃ and 10 min activation period at 95 ℃ followed by 45 cycles of a two-step thermal profile of 30 s at 95 ℃ denaturation and 30 s at 54℃/60 ℃ for combined annealing-extension and 1 cycle of 60 ℃ for 1 min. 4 μL template was added to the 30 μL reaction system (one-step) of the BioDigital Yan platform, the thermal cycling consisted of a 10 min reverse transcription at 50℃ and 10 min activation period at 95 ℃ followed by 45 cycles of a two-step thermal profile of 20 s at 95 ℃ denaturation and 40 s at 54℃/60 ℃ for combined annealing-extension and 1 cycle of 25 ℃ for 10 s. 4 μL template was added to the 40 μL reaction system (one-step) of the QIAcuity One platform, the thermal cycling consisted of a 40 min reverse transcription at 50℃ and 2 min activation period at 95 ℃ followed by 40 cycles of a two-step thermal profile of 5 s at 95 ℃ denaturation and 30 s at 54℃/60 ℃.

Homogeneity assessment

Eleven vials were randomly selected from the batch, and each was analyzed in three replicates. Measurements were performed by quantifying the copy number of *EML*4-*ALK* fusion variant and *ALK*-ref and the ratio of *EML*4-*ALK*/*ALK* using the established RT-dPCR method. The Shapiro-Wilke method was first used to check the normality of the data, and the homogeneity data of all RMs satisfied the normal distribution. Dixon's and Grubbs's tests showed no suspicious values. All test results of their *F*-test for RMs were lower than the *F* critical value. That is, there is no significant difference in the internal homogeneity of the reference material solution (Supplemental Tables 8 and 9).

Stability assessment

Stability of the produced batch of materials was evaluated using a classical stability study design that followed ISO guide 35.

Short-term stability is to ensure the stability of the RMs under the transportation conditions, so the

two temperatures of -20 ℃ and 4 ℃ were conducted, and 3 vials of samples were randomly selected at the sampling time points of 0 days, 3 days, 7 days and 14 days, respectively, and each was analyzed in three replicates to quantifying the copy number of each *EML*4-*ALK* fusion variant, *ALK*-ref and the ratio of *EML*4-*ALK*/*ALK*. The results were shown in Supplemental Figures 7 and 8, and there was no significant change in each *EML*4-*ALK* fusion variant, *ALK*-ref and the ratio of *EML*4-*ALK*/*ALK* within 14 days when the gRNA reference material was stored at -20 ℃ and 4 ℃.

Long-term stability at -80 °C was continuously monitored for 12 months at the time of this study. The *EML*4-*ALK* gRNA RMs was investigated for long-term stability in 0, 1, 2, 3, 6 and 12 months. The slope(b1) of the regression line of the measured values at each storage time and the standard deviation of the standard deviation of slope（$s(b_1)$）were calculated for the two RMs. Two of *EML*4-*ALK* fusion variants RMs were stable since the $|b_1| < s(b_1)*t_{0.95,n-2}$, whereas $t_{0.95,n-2}$ is a t-factor based on n-2 degrees of freedom for the tested time point at 95% confidence level (Supplemental Tables 10 and 11).

Measurement uncertainty of the reference material value

The combined uncertainty of the property value ($u_{c,rel}$) containing uncertainty of method characterization ($u_{char,rel}$), homogeneity ($u_{bb,\ rel}$) and stability ($u_{s,rel}$) is shown in Supplemental Tables 12 and 13. The measurement uncertainty associated with batch characterization includes several aspects related to the procedure used to quantify the copy number of each *EML*4-*ALK* fusion variant, *ALK*-ref and the ratio of *EML*4-*ALK*/*ALK*. This parameter includes the uncertainty related to the precision of the RT-dPCR method ($u_{a,\ rel}$), as well as the uncertainty related to the type B uncertainty of pipette aspiration volume and droplet volume, as previously described[1,2]. The uncertainty caused by stability was calculated by the standard deviation of the regression coefficient ($s(b_1)$) multiplying the storage time (taken for 12 months) according to ISO guideline 35. The uncertainty of homogeneity accounts for the differences between units. The relative standard uncertainty of the above three parts was combined to obtain the combined standard uncertainty ($u_{c,rel}$) of the property value of the RM according to equation (3). Taking a coverage factor of 2, the extended uncertainty (U_c) of the reference material can be obtained at a 95% confidence level according to the formula Uc=k×u_c. The expanded uncertainty for the ratio of *EML*4-*ALK*/*ALK* were 12% and 8.0%, respectively.

$$u_c = \sqrt{u_{char}{}^2 + u_{bb}{}^2 + u_s{}^2} \qquad (3)$$

Reference

1 Dong, L. *et al.* Comparison of four digital PCR platforms for accurate quantification of DNA

copy number of a certified plasmid DNA reference material. *Sci Rep.* **5**, 13174. doi:10.1038/srep13174 (2015).

2 Niu C, *et al.* Research on the Collaborative Value Assignment of Plasmid Nucleic Acid Reference Material Based on Digital PCR. *Acta Metrologica Sinica*, **42**(11), 1522-1527 (2021).

Supplemental Tables and Figures

Supplemental Table 1 Primer and probe sequences used in the study

Target	Sequence（5'-3'）	Reference
	EML4-ALK	
V1-F	GCCCACACCTGGGAAAGG	
V3-F	AAACTGCAGACAAGCATAAAGATGTC	[3]
ALK-Exon20-R	GGAGCTTGCTCAGCTTGTACTCA	
ALK-Exon20-P	FAM-CCGCCGGAAGCACCAGGAGC-BHQ1	
	ALK-ref	
ALK -F	GGCAAGACCTCCTCCATCAG	This study
ALK -R	GCCCAGACCCCGAATGAG	
ALK- -P	HEX-AGGTGCCGCGGAAA-BHQ1	

Supplemental Table 2 Summary of the RT-dPCR platforms and methods of laboratories

ID	Platform	Approach	RT-PCR (Mix)	Reaction/ Sample volume	Thermal cycling conditions
1	QX200	One-Step	One-step RT-ddPCR Kit for Probes(Bio-Rad)	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min
2	QX200	One-Step	One-step RT-ddPCR Kit for Probes(Bio-Rad)	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min
3	QX200	One-Step	One-step RT-ddPCR Kit for	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，

			Probes(Bio-Rad)		30s, 54℃/60℃，1 min);98℃，10 min
4	TD-1	Two-Step	RT-ddPCR Mixtures	30μL/2μL	55℃，15min;95℃，15 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min
5	DQ24	One-Step	One-step RT-ddPCR Kit for Probes(Sniper)	22μL/2μL	50℃，10 min;95℃，15 min;40 cycles of (95℃，15s, 54℃/60℃，30 s);60℃，1 min
6	QX200	One-Step	One-step RT-ddPCR Kit for Probes(Bio-Rad)	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min
7	QX200	One-Step	One-step RT-ddPCR Kit for Probes(Bio-Rad)	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min
8	QX200	One-Step	One-step RT-ddPCR Kit for Probes(Bio-Rad)	20μL/2μL	45℃，60 min;95℃，10 min;40 cycles of (94℃，30s, 54℃/60℃，1 min);98℃，10 min

Supplemental Table 3 Correlation of *EML4-ALK* V1 measurement between gravimetrical dilution and RT-dPCR.

Gravimetric	RT-dPCR	CV (%)
120650	125467	3.47
15194	15253	2.35
1800	1744	5.59
180	178	9.37
37	36	24.56
5	9	55.78
2	3	64.27

Supplemental Table 4 Correlation of *EML4-ALK* V3 measurement between gravimetrical dilution and RT-dPCR.

Gravimetric	RT-dPCR	CV (%)
120630	120933	5.79
15069	15980	5.97
1247	1205	7.11
123	126	12.52
49	48	24.40
6	7	82.85
2	1	95.80

Supplemental Table 5 Correlation of *ALK*-ref measurement between gravimetrical dilution and RT-dPCR.

Gravimetric	RT-dPCR	CV (%)
120630	117600	1.67
13482	13482	0.33
1371	1362	1.85
134	129	4.25
13	15	14.21
7	8	26.46

1	2	62.92

Supplemental Table 6 The 15 highest values of the LoB (copies/reaction)

No.	V1	V3	*ALK*-ref
60	1.90	1.20	0.50
59	0.90	0.90	0.50
58	0.84	0.90	0.20
57	0.84	0.90	0.20
56	0.80	0.74	0.20
55	0.80	0.74	0.20
54	0.80	0.70	0.00
53	0.80	0.64	0.00
52	0.70	0.60	0.00
51	0.50	0.60	0.00
50	0.42	0.60	0.00
49	0.42	0.60	0.00
48	0.30	0.50	0.00
47	0.30	0.50	0.00
46	0.20	0.50	0.00

Supplemental Table 7 The RTE assessment results for V1, V3, and *ALK*-ref

Type	V1	V3	*ALK*-ref
RT-dPCR (copies μL^{-1})	1.16E+09	1.03E+09	1.08E+09
ID-MS (copies μL^{-1})		1.08E+09	
RTE	1.07	0.949	0.995

Supplemental Table 8 Homogeneity assessment of *EML4-ALK* V1

Units	**V1（copies/μL）**			***ALK*-ref（copies/μL）**			***EML*4-*ALK*/*ALK*（%）**		
1	3651	3681	3791	4480	4490	4430	81.5	82.0	85.6
2	3671	3661	3761	4480	4490	4430	81.9	81.5	84.9
3	3861	3861	3611	4510	4450	4580	85.6	86.8	78.8
4	3741	3701	3711	4510	4450	4580	82.9	83.2	81.0
5	3541	3631	3711	4520	4510	4600	78.3	80.5	80.7

6	3741	3501	3641	4550	4460	4460	82.2	78.5	81.6
7	3751	3621	3871	4410	4420	4420	85.0	81.9	87.6
8	3781	3661	3761	4510	4360	4470	83.8	84.0	84.1
9	3641	3581	3711	4430	4480	4560	82.2	79.9	81.4
10	3631	3731	3681	4420	4430	4480	82.1	84.2	82.2
11	3581	3601	3711	4500	4470	4430	79.6	80.5	83.8
Q_1		82588			40151			68.6	
V_1		10			10			10	
S_1^2		8259			4015			6.9	
Q_2		165800			56600			100.1	
V_2		22			22			22	
S_2^2		7536			2572			4.5	
F		1.1			1.7			1.5	
$F_{0.05(10,22)}$		2.3			2.3			2.3	
CV (%)		1.0			0.9			1.8	
u_{bbrel}(%)		0.7			0.4			0.8	

Supplemental Table 9 Homogeneity assessment of *EML4-ALK* V3

Units	**V3（copies/μL）**			***ALK*-ref（copies/μL）**			***EML*4-*ALK*/*ALK*（%）**		
1	13297	14017	13937	13990	13950	13970	95.0	100.5	99.8
2	14047	12927	13067	13990	13920	13940	100.4	92.9	93.7
3	13907	13667	13387	13970	13980	13900	99.5	97.8	96.3
4	13367	13627	13097	13980	14000	14000	95.6	97.3	93.6
5	13867	13657	14037	13960	14010	13910	99.3	97.5	100.9
6	13567	13627	13747	14060	13970	14070	96.5	97.5	97.7
7	13217	13487	13197	14020	13970	14000	94.3	96.5	94.3
8	13797	13447	13997	14010	13910	13950	98.5	96.7	100.3
9	13757	13877	13377	13970	13980	14120	98.5	99.3	94.7
10	13327	13577	13387	14040	14010	14120	94.9	96.9	94.8

11	13027	13747	13477	14020	14030	13910	92.9	98.0	96.9
Q_1	1123563			39284			64.5		
V_1	10			10			10		
S_1^2	112356			3928			6.5		
Q_2	2063867			54266			109.5		
V_2	10			22			22		
S_2^2	93812			2467			5.0		
F	1.2			1.6			1.3		
$F_{0.05(10,22)}$	2.3			2.3			2.3		
CV (%)	1.5			0.3			1.5		
u_{bbrel}(%)	0.7			0.1			0.7		

Supplemental Table 10 Result of long-term stability study for *EML4-ALK* V1

Time	**V1 (copies/µL)**	***ALK*-ref (copies/µL)**	***EML*4-*ALK*/*ALK* (%)**
0	3592	4167	86.3
1	3345	4118	81.4
2	3730	4428	84.0
3	3494	4381	80.2
6	3581	4427	81.4
12	3712	4108	90.8
$\lvert b1 \rvert$	16	6	0.6
$s(b_1)$	15	18	0.4
$s_{(b1)} * t_{0.95,5}$	40	48	1.0
Conclusion	Stable	Stable	Stable

Supplemental Table 11 Result of long-term stability study for *EML4-ALK* V3

Time	**V3 (copies/µL)**	***ALK*-ref (copies/µL)**	***EML*4-*ALK*/*ALK* (%)**
0	13382	13940	94.6
1	13967	14017	99.6

2	12191	12622	96.5
3	12952	14022	96.8
6	12922	13388	96.6
12	13462	13404	100.4
$\lvert b1 \rvert$	8	32	0.3
$s(b_1)$	68	60	0.2
$s_{(b1)} * t_{0.95,5}$	189	166	2.4
Conclusion	Stable	Stable	Stable

Supplemental Table 12 Reference value, factors contributing to the relative standard uncertainty and expanded uncertainty of the *EML4-ALK* V1 reference material

Property	**V1 (copies/µL)**	***ALK*-ref (copies/µL)**	***EML*4-*ALK*/*ALK* (%)**
Reference value	3395	4512	74.9
$U_{char,rel}$ (%)	2.0	1.1	1.5
$u_{bb\ rel}$ (%)	0.7	0.4	0.8
$u_{S\ rel}$ (%)	4.7	4.6	5.4
$u_{RTE,rel}$(%)	10.0	10.0	-
$u_{c\ rel}$ (%)	11.3	11.1	5.7
U_{rel} (k=2) (%)	23	23	12

Supplemental Table 13 Reference value, factors contributing to the relative standard uncertainty and expanded uncertainty of the *EML4-ALK* V3 reference material

Property	**V3 (copies/µL)**	***ALK*-ref (copies/µL)**	***EML*4-*ALK*/*ALK* (%)**
Reference value	13804	13582	101.6
$U_{char,rel}$ (%)	2.5	2.4	2.5
$u_{bb\ rel}$ (%)	0.7	0.1	0.7
$u_{S\ rel}$ (%)	6.0	5.1	2.4
$u_{RTE,rel}$(%)	10.0	10.0	-
$u_{c\ rel}$ (%)	12.0	11.5	3.6

U_{rel} (k=2) (%)	24	23	8.0

u_{char}, uncertainty of characterization by RT-dPCR;

u_{bb}, uncertainty of homogeneity;

u_S, uncertainty of the long-term stability (12 months);

$u_{RTE,}$ uncertainty of the reverse transcription efficiency (RTE);

$u_{c,rel}$, the combined uncertainty of the above four components.

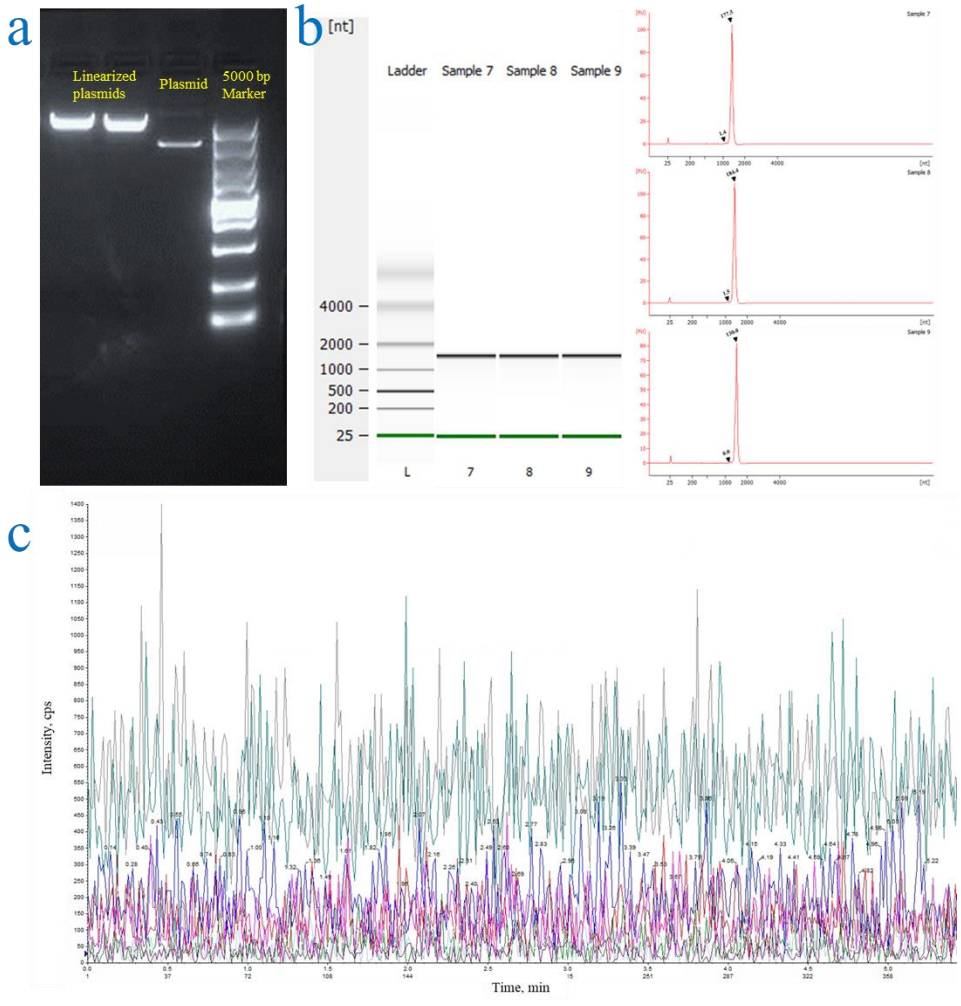
Supplemental Table 14 Interlaboratory result of RM1 for reverse transcription efficiency

Lab	**V1 (copies/μL)**		***ALK*-ref (copies/μL)**		***EML*4-*ALK*/*ALK* (%)**	
1	3052	111	4435	62	68.8	4.9
2	3306	69	4384	162	75.4	4.4
3	3326	165	4605	59	72.2	3.6
4	3349	64	4412	73	75.9	0.7
5	3606	58	4795	251	75.2	5.7
6	3669	148	4657	138	78.8	1.7
7	3487	53	4496	91	77.5	1.9
8	3365	124	4494	25	74.9	3
Average	3395		4535		74.8	
SD	193		141		3.1	
CV(%)	5.7		3.1		4.1	

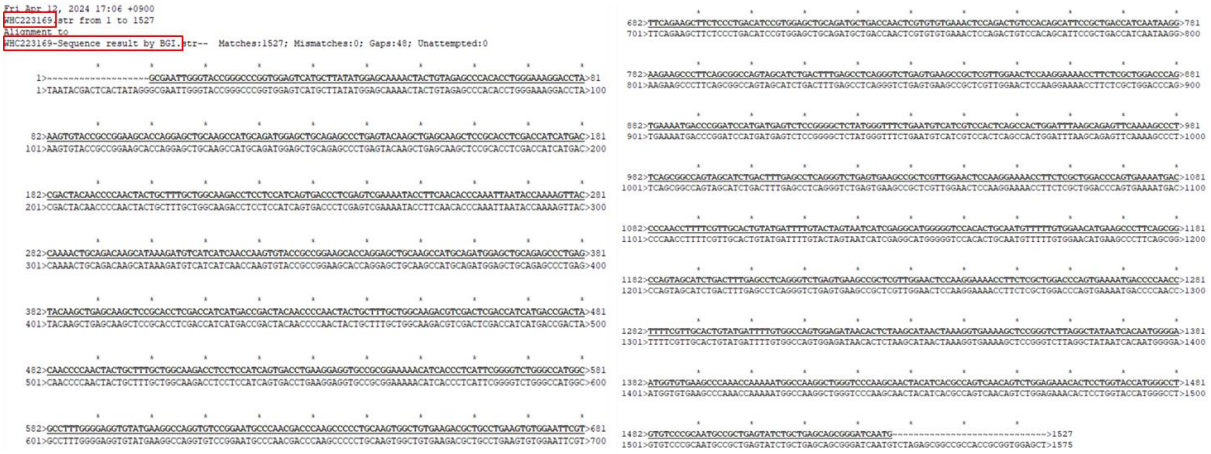
Supplemental Table 15 Interlaboratory result of RM2 corrected for reverse transcription efficiency

Lab	**V3 (copies/μL)**		***ALK*-ref (copies/μL)**		***EML*4-*ALK*/*ALK* (%)**	
1	11623	99	13148	68	88.4	8.4
2	13621	101	12989	110	104.9	1.6
3	14033	66	13493	122	104.0	1
4	13793	148	12104	56	114.0	1.6
5	14731	252	15155	243	97.2	2.8
6	14745	93	14158	81	104.1	1.2

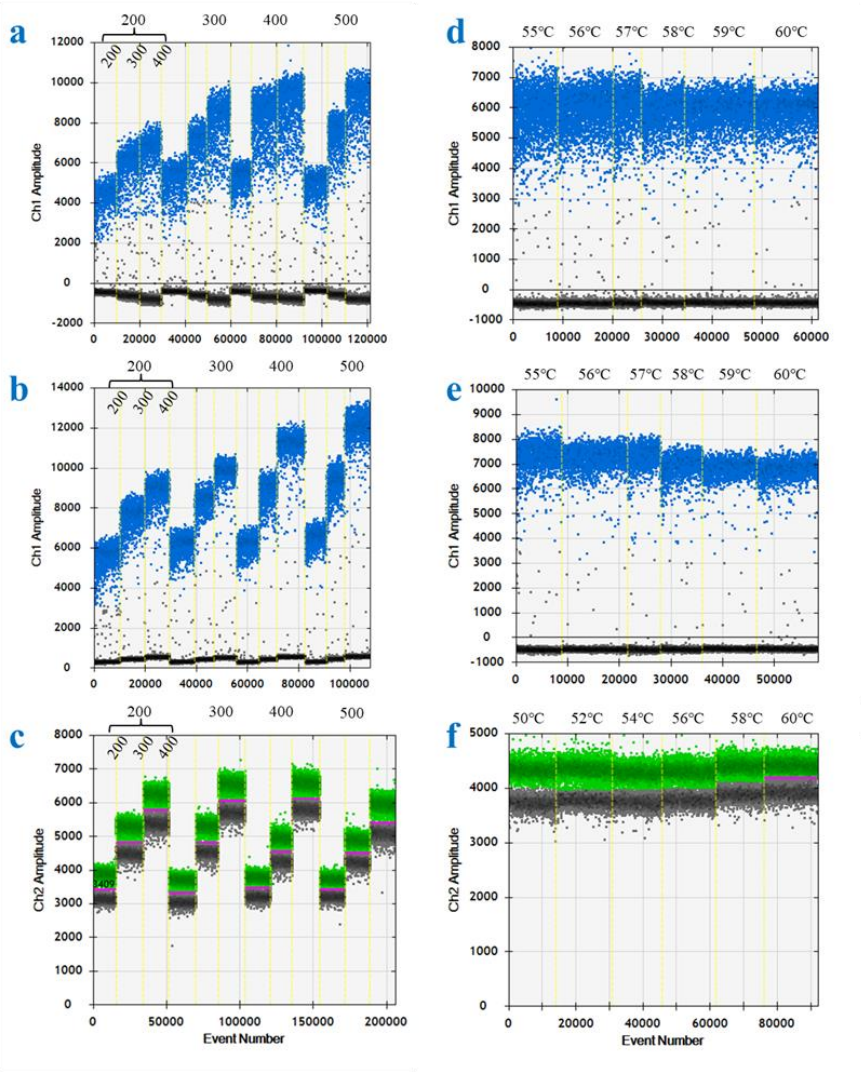
7	13573	167	13527	16	100.3	1.1
8	14316	54	14082	51	101.7	1.4
Average	13804		13582		101.8	
SD	993		910		7.3	
CV (%)	7.2		6.7		7.1	



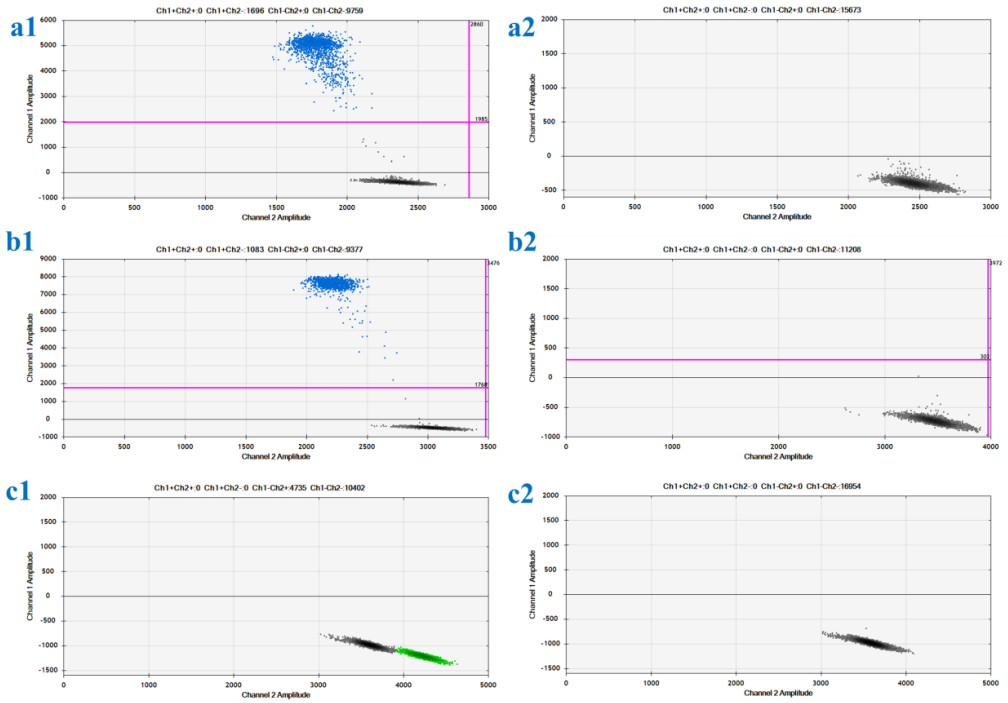
Supplemental Figure 1. Length and purity validation results during in vitro transcribed RNA preparation. a, the linearized plasmid (4402 bp) was checked using 2% agar gel electrophoresis; b, in vitro transcribed RNA (1527 nt) array electrophoresis results by Bioanalyzer (L: ladder, sample 7-9: in vitro transcribed RNA, the number annotated is the peak area.); c, the result of undigested in vitro transcription RNA by mass spectrometry.



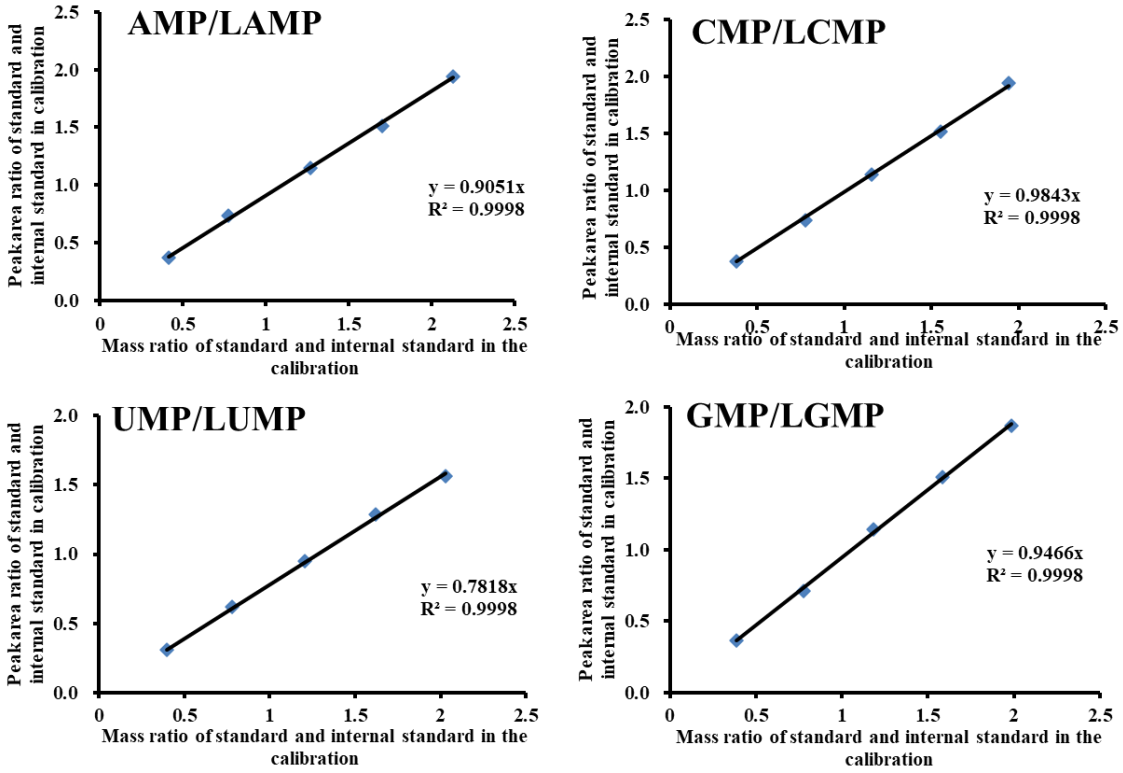
Supplemental Figure 2. In vitro transcribed RNA sequencing alignment results.



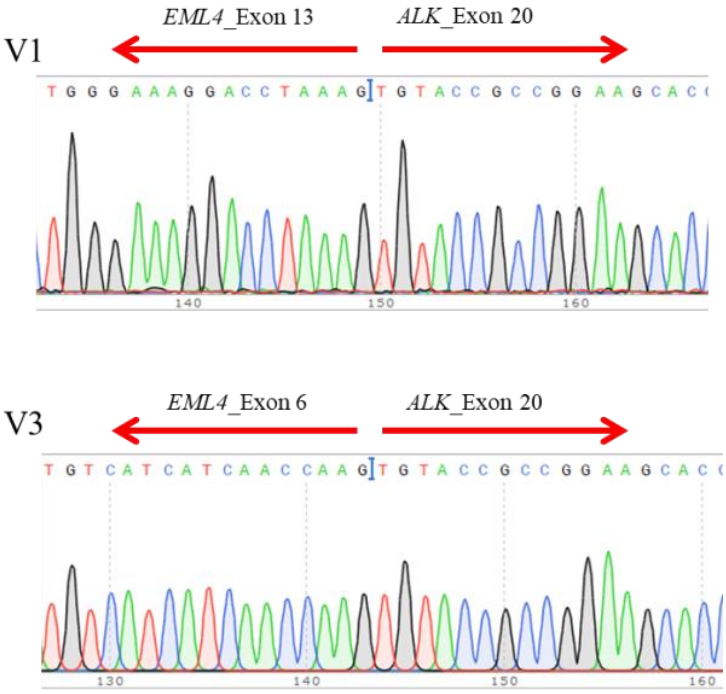
Supplemental Figure 3. Optimization of annealing temperature and primer and probe working concentration of RT- dPCR for *EML4-ALK* fusion variant and *ALK*-ref. a-c, optimization of primer and probe working concentration for V1, V3 and *ALK*-ref, respectively. d-f, optimization of annealing temperature for V1, V3 and *ALK*-ref, respectively.



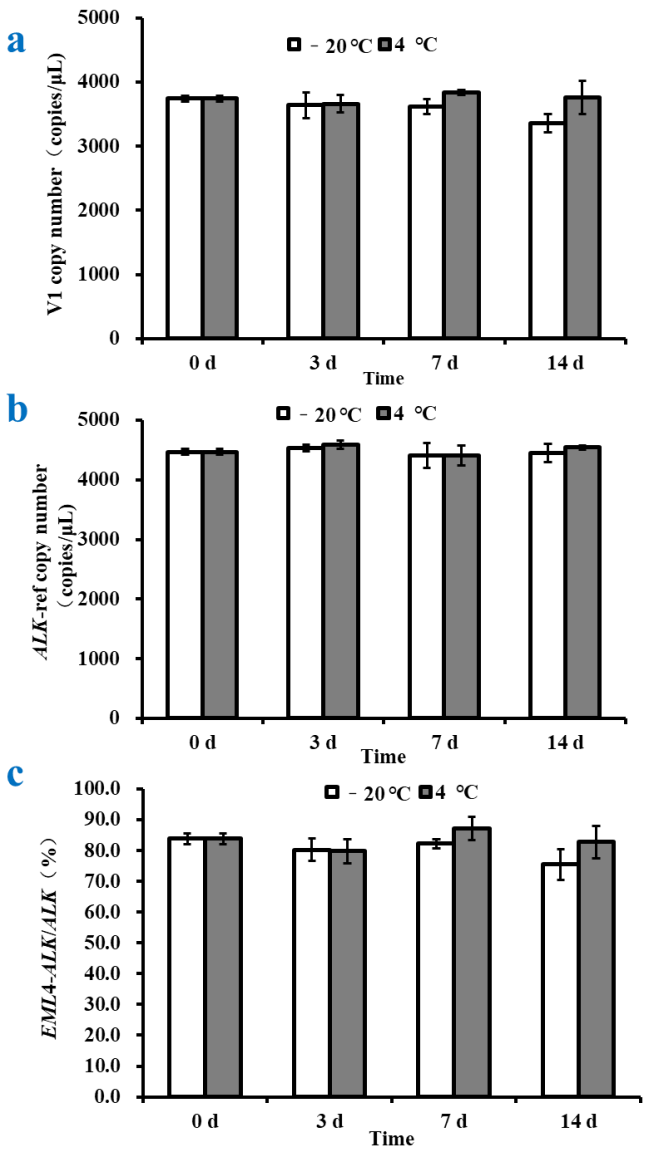
Supplemental Figure 4. Two-dimensional scatter plot for selected wells of *EML4-ALK* fusion variant and *ALK*-ref. a1, V1 + in vitro transcribed V1, a2, no template control for V1, b1, V3 + in vitro transcribed V3, b2, no template control for V3, c1, *ALK*-ref + in vitro transcribed *ALK*, c2, no template control for *ALK*.



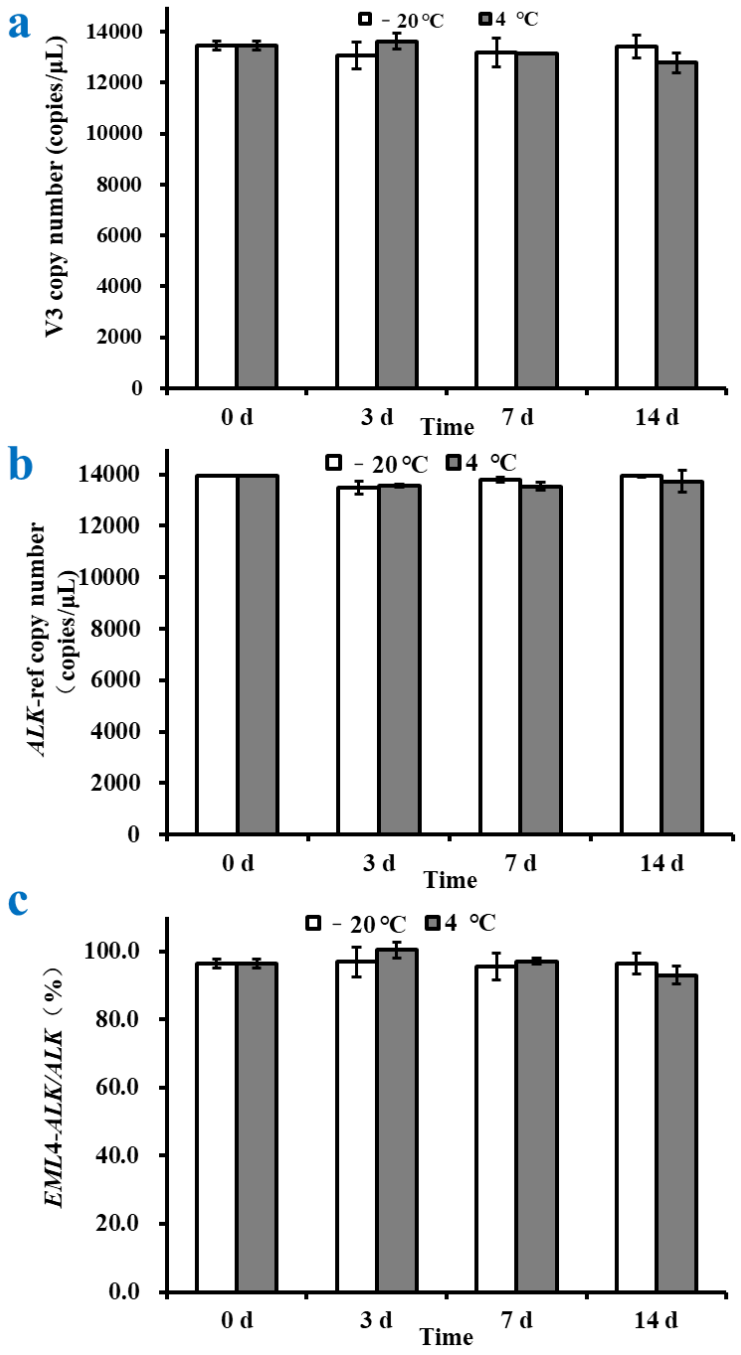
Supplemental Figure 5. Calculation curve of four NMPs.



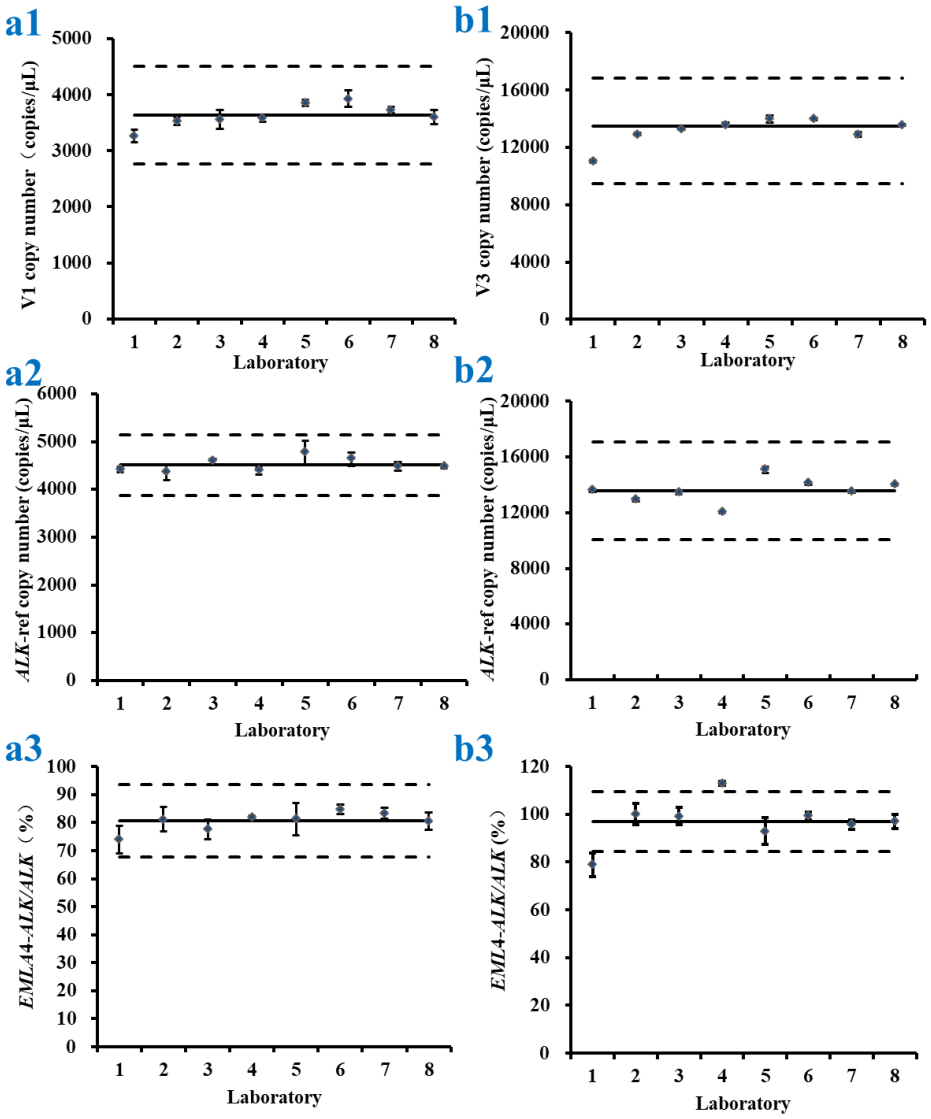
Supplemental Figure 6. Evidence of *EML4-ALK* fusion variants in the two candidate reference materials confirmed by Sanger sequencing



Supplemental Figure 7. Short-term stability data of the copy number of V1 (a) and *ALK*-ref (b) and the ratio of *EML4-ALK*/*ALK* (c) in *EML4-ALK* V1 reference materials



Supplemental Figure 8. Short-term stability data of the copy number of V3 (a) and *ALK*-ref (*b*) and the ratio of *EML*4-*ALK*/*ALK* (c) in *EML*4-*ALK* V3 reference materials



Supplemental Figure 9. Interlaboratory assessment of candidate reference measurement procedure of *EML*4-*ALK* V1 reference material (a) and *EML*4-*ALK* V3 reference material (b).