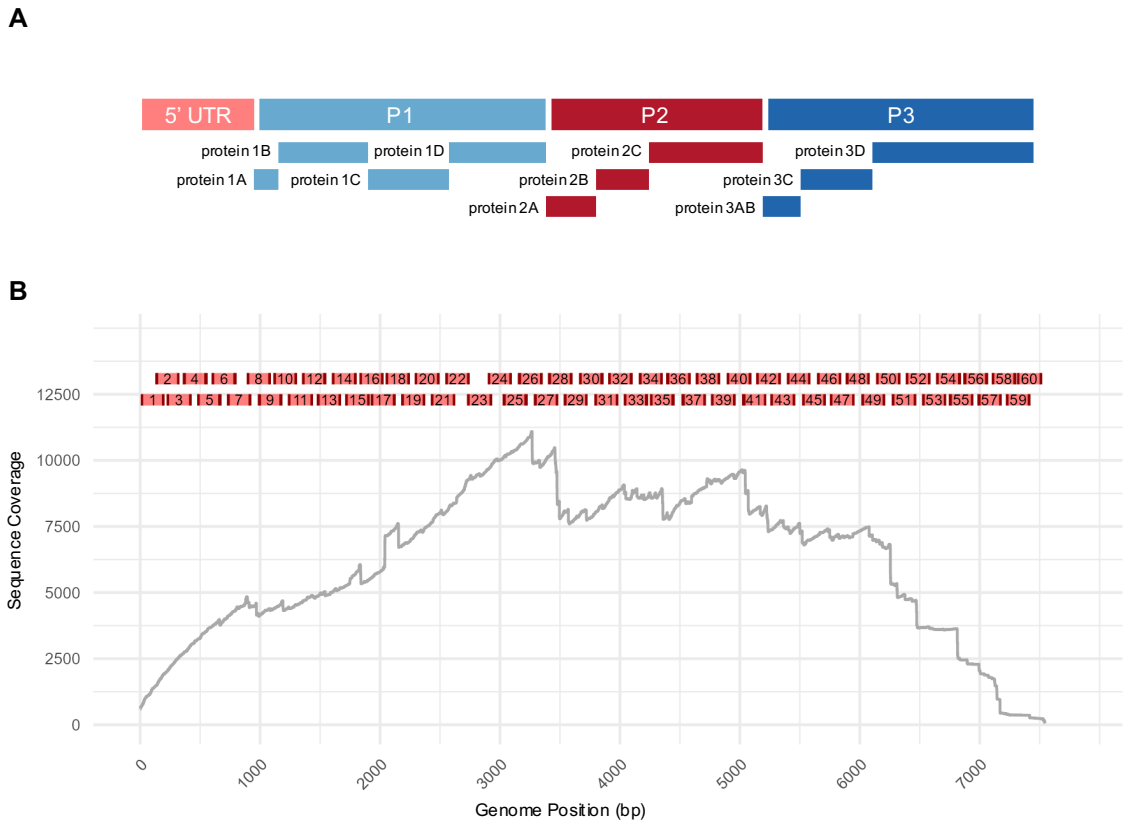
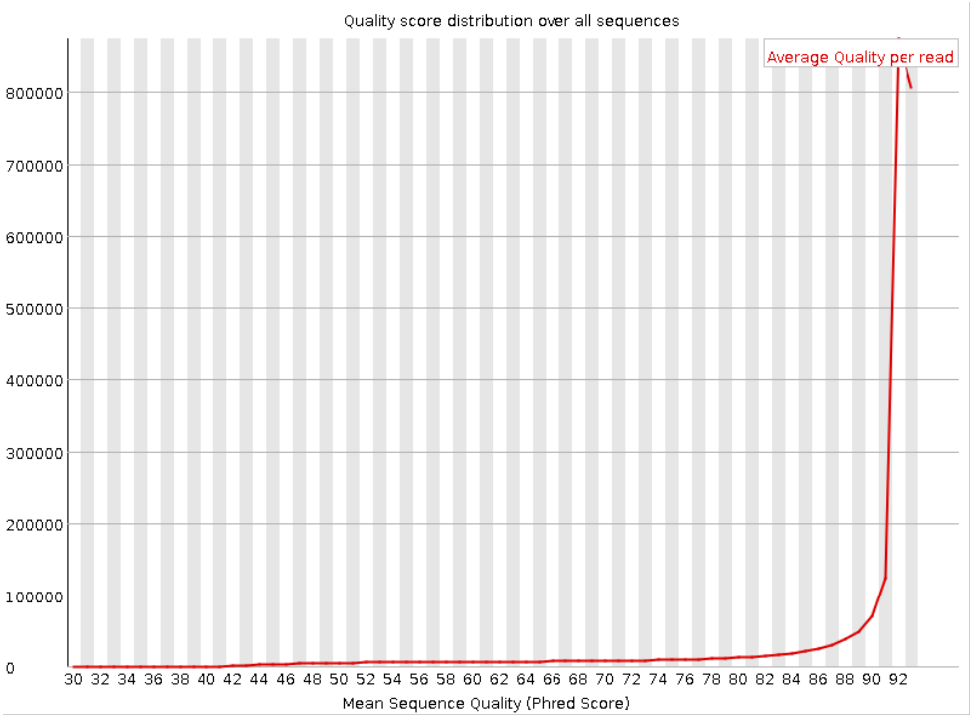
SUPPLEMENTARY MATERIAL

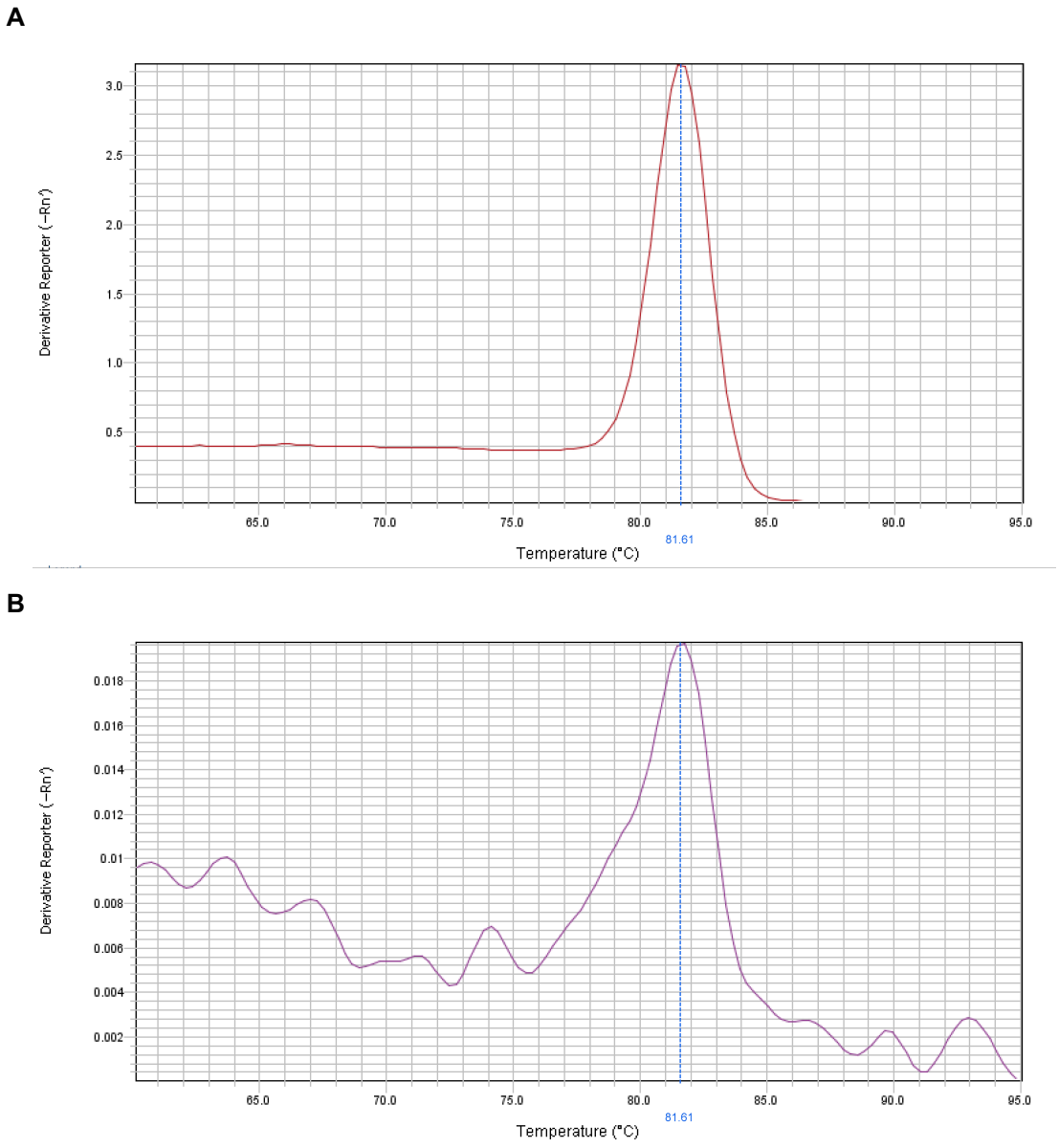
Supplementary_Table_S1	**Catalogue of the Mengovirus amplicon sequences. It includes sequence names, genomic coordinates, orientation, nucleotide sequences, sequence length, GC content, and melting temperature.**
Supplementary_Table_S2	**Analysis of amplicon metrics across various regions**. It details predicted melting temperatures, alongside qPCR-derived Ct values and sequencing coverage statistics.
Supplementary_Table_S3	**Compiled data from the top three primers, selected from an initial set of 60, based on their performance in amplifying RNA and library HIV samples.** It presents detailed amplification metrics, including Ct values and ΔCt calculations, providing a comparative evaluation of primer efficacy under different conditions.
Supplementary_Table_S4	**Efficiency of HIV and Mengovirus primers across various extraction conditions.** Focusing on qPCR results (Ct values) and sequencing output, it includes primer sequences, amplification efficiency, correlation coefficients (R^2), slopes.
Supplementary_Table_S5	**Clinical viral loads with corresponding qPCR and sequencing results.** It includes clinical viral load measurements, Ct values from different primer sets, and calculated viral loads based on RNA and Library.



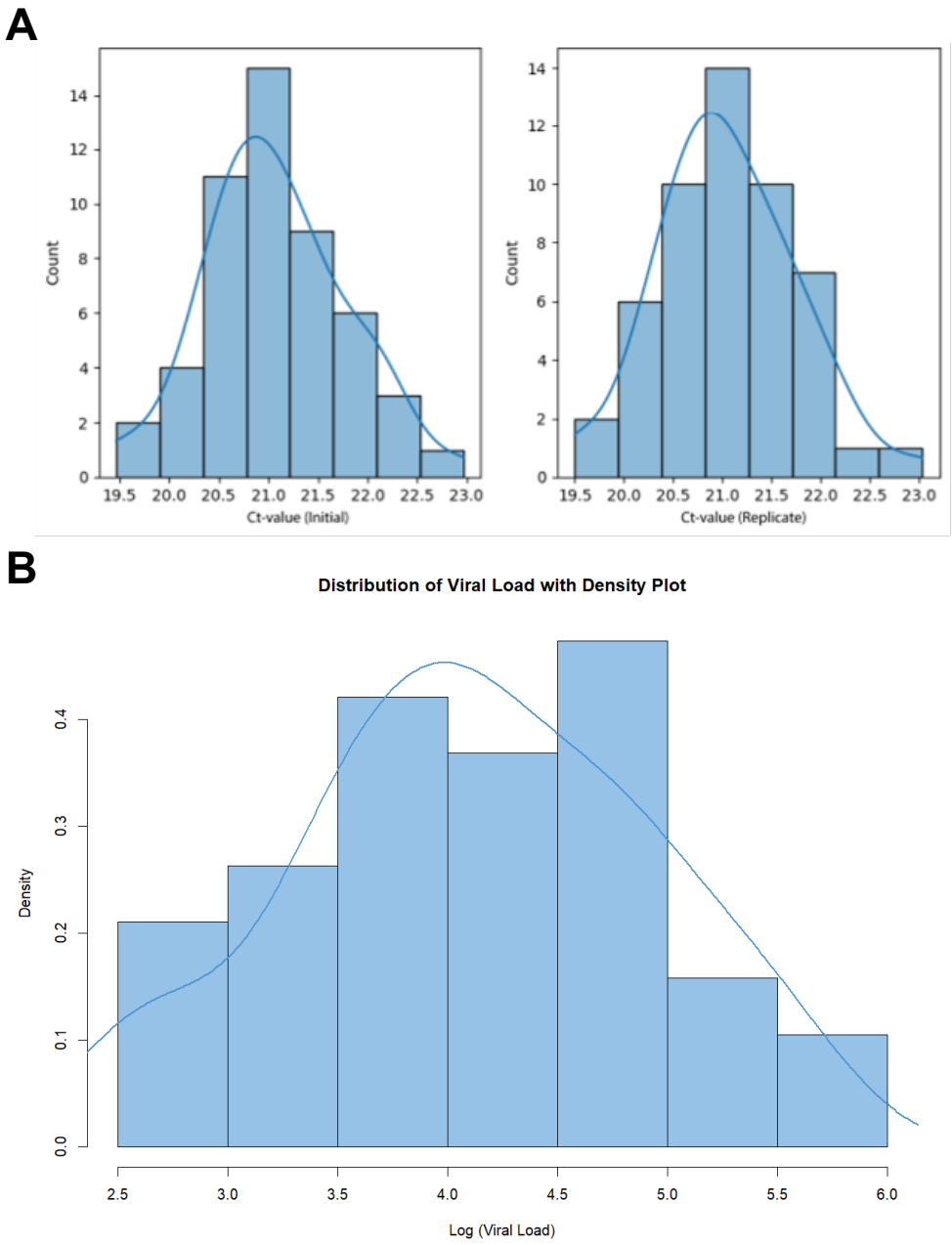
Supplementary Figure S1: Mengovirus genome structural overview and sequencing coverage. (A) The Mengovirus genome is divided into three regions: P1 (capsid proteins), P2, and P3 (non-structural proteins for replication). This 7,597 nucleotide-long genome is a single-stranded positive-sense RNA with a 5' viral genome protein (VG) and a 3' polyadenine tail, translated directly by host ribosomes and cleaved into functional proteins by viral proteases (74, 78, 79). **(B)** Sequencing coverage across the Mengovirus genome. The x-axis represents genome position (bp), and the y-axis shows sequencing coverage depth. The line graph illustrates the sequencing coverage depth at each position along the genome. Primer set labels (red) are aligned above the plot according to their corresponding positions, facilitating the correlation between primer design and sequencing coverage. Coverage is notably lower at the 5' UTR and increases towards the middle regions.



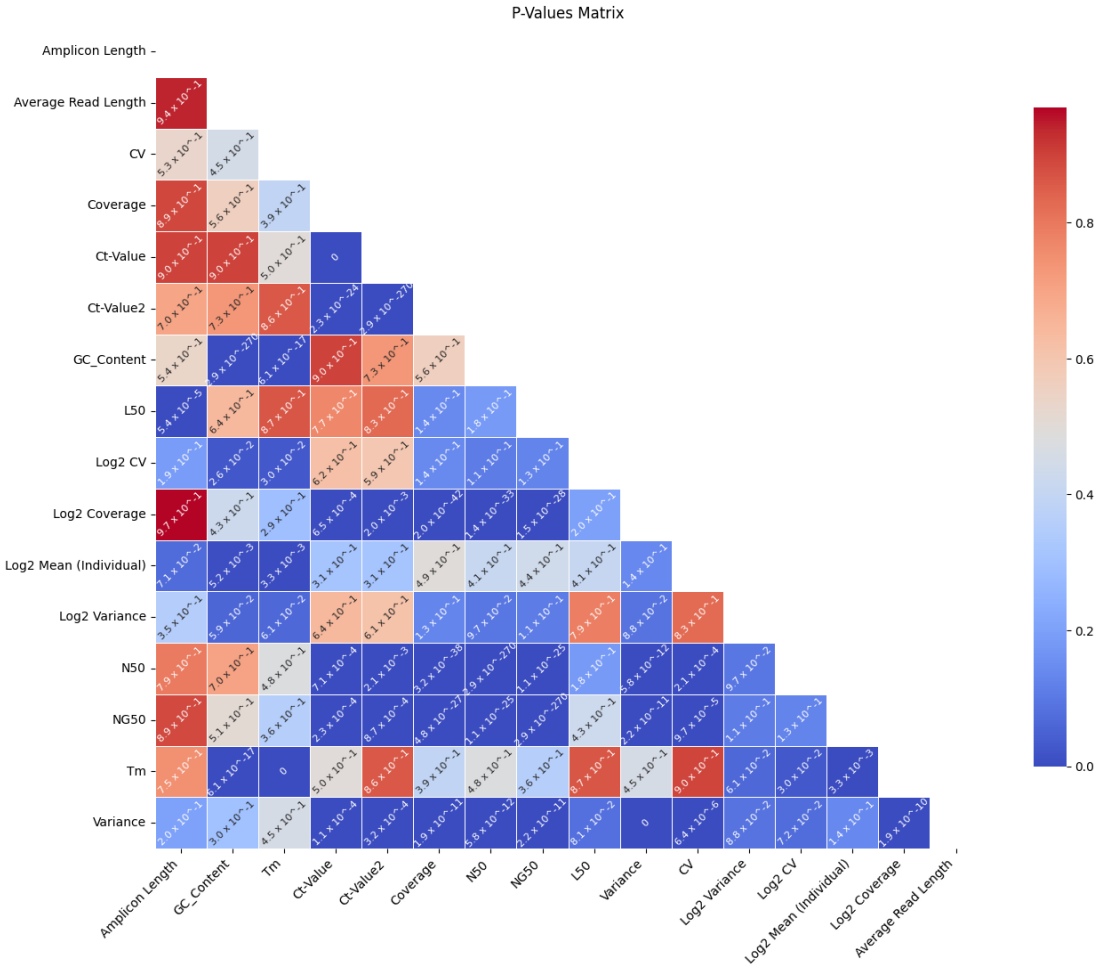
Supplementary Figure S2: FastQC Quality Metrics for Mengovirus Sequencing Data. The quality score distribution over all sequences shows that most reads exhibit high Phred scores.



Supplementary Figure S3: Melt Curve Analysis of qPCR Products for Mengovirus Detection. (A) Primer Pair 15: melt curve analysis of a qPCR product demonstrating a single sharp peak at 81.61°C indicates a specific and high-quality amplification product. The derivative reporter (-dRn/dT) is plotted against temperature (°C), showing a well-defined melt peak. It helps identify the temperature at which the dsDNA template denatures (melts). Peaks in the derivative plot correspond to the amplified products' melting temperatures (Tm). **(B) Primer Pair 6**: melt curve analysis of a qPCR product showing multiple peaks with the prominent peak at 81.61°C but of low amplitude. The presence of additional peaks indicates potential non-specific amplification or primer-dimer formation. The derivative reporter (-dRn/dT) is plotted against temperature (°C), highlighting the variability in the melt profile. Complete data is available in **S_Table2**.



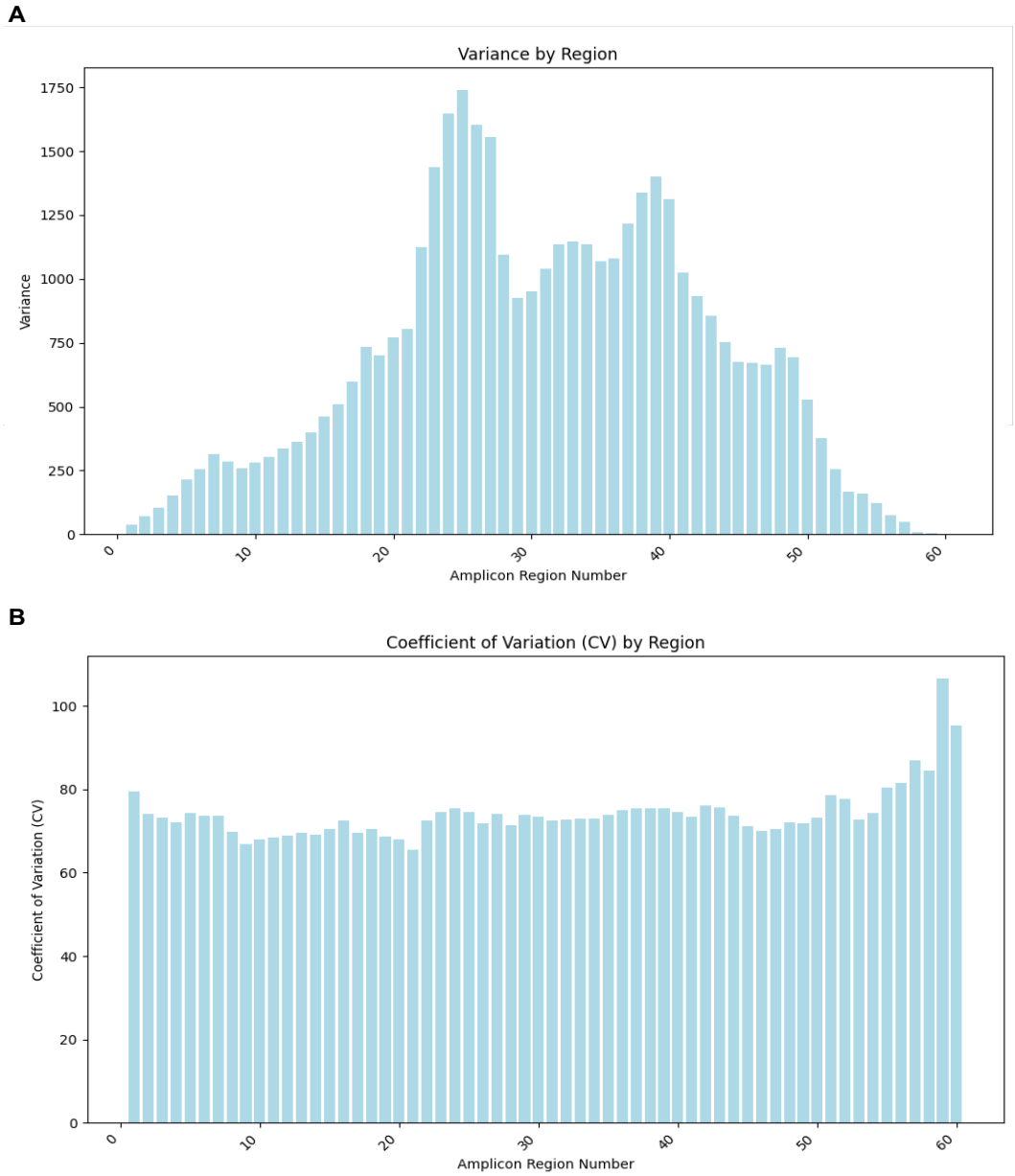
Supplementary Figure S4: (A) Distribution of Ct value Frequency Across 60 Amplicons. The bars represent 1/8th of Ct value range intervals. The line represents a normal distribution fit to the data. Note the similar distribution of initial and technical replicate Ct values. **(B)** Viral Loads Across 38 Samples: The bars represent 1/7th intervals of the viral load (log-transformed) range of 38 HIV clinical samples. The line represents a normal distribution fit to the data.



Supplementary Figure S5: P-Values Matrix of 60 Amplicon Sequences between various experimental parameters, including amplicon length, GC content, coverage, and Ct values.

Supplementary Table S6: P-values for the pairwise correlations between amplicon features and sequencing metrics. Values below 0.05 are bold.

	Amplicon Length	**GC Content**	**Tm**	**Ct Value**	**Coverage**	**N50**	**LG50**	**Coverage Variance**	**Coverage CV**	**Average Read Length**
Amplicon Length	**0**	8.09E-01	5.45E-01	5.86E-01	3.57E-01	3.54E-01	4.30E-01	9.59E-01	2.51E-01	5.59E-01
GC Content	8.09E-01	**0**	**8.87E-26**	5.39E-01	4.46E-01	5.07E-01	2.37E-01	9.55E-01	2.55E-01	9.85E-01
Tm	5.45E-01	**8.87E-26**	**0**	8.47E-01	3.69E-01	4.13E-01	5.53E-02	8.49E-01	3.59E-01	4.89E-01
Ct Value	5.86E-01	5.39E-01	8.47E-01	**0**	7.93E-01	7.55E-01	1.05E-01	8.74E-01	**6.29E-03**	4.34E-01
Coverage	3.57E-01	4.46E-01	3.69E-01	7.93E-01	**0**	**1.65E-69**	**2.61E-06**	**3.26E-21**	**4.41E-06**	**6.72E-08**
N50	3.54E-01	5.07E-01	4.13E-01	7.55E-01	**1.65E-69**	**0**	**4.06E-06**	**1.40E-21**	**6.32E-06**	**1.37E-07**
LG50	4.30E-01	2.37E-01	5.53E-02	1.05E-01	**2.61E-06**	**4.06E-06**	**0**	**5.69E-04**	**5.33E-03**	**2.57E-02**
Coverage Variance	9.59E-01	9.55E-01	8.49E-01	8.74E-01	**3.26E-21**	**1.40E-21**	**5.69E-04**	**0**	**3.37E-02**	**3.29E-05**
Coverage CV	2.51E-01	2.55E-01	3.59E-01	**6.29E-03**	**4.41E-06**	**6.32E-06**	**5.33E-03**	**3.37E-02**	**0**	**1.18E-03**
Average Read Length	5.59E-01	9.85E-01	4.89E-01	4.34E-01	**6.72E-08**	**1.37E-07**	**2.57E-02**	**3.29E-05**	**1.18E-03**	**0**



Supplementary Figure S6: Variance and Coefficient of Variation (CV) by Amplicon Region. (A) Variance by region shows the distribution of variances for each amplicon region across 57 samples with equal quantities of Mengo virus control spiked into each sample. The variance mirrors sequencing coverage, indicating that areas with higher sequencing depth tend to have higher variance. **(B)** The coefficient of variation (CV) by region displays the relative variability across different amplicon regions, measuring consistency in the context of each amplicon region.

Supplementary Table S7: Discarded Primers due to Unspecific Amplification.

Primer Pair	Target Region	Forward Primer	Reverse Primer	Ct values (HIV-free) *	Ct values (HIV + Mengovirus) *
Primer Pair 03	5' UTR	GACAAGCAACGTCTGTAGCGA	TACCTTCTGGGCATCCTTCAGC	21.282/21.203	20.931/20.896
Primer Pair 30	P2 (Protein 2A)	CTTTATAATCTTCATTGCTTTTTGGGCG	GCAAAGCTATGGGATCAGCCTT	20.748/20.857	20.525/20.462
Primer Pair 43	P3 (Protein 3AB)	GGATGAAGATTTCCGCCATGCT	TCCATAGTCGGATTTGGACCCT	21.032/21.075	20.771/20.726
Primer Pair 49	P3 (Protein 3C)	CACAAGAAATGATCGATGCGGTG	ACTTCGTCAACATCAGCATCCG	20.577/20.545	20.252/20.243
Primer Pair 53	P3 (Protein 3D)	CAGACATTCCTCAAGGACGAGC	AAAGGCTGTCCAATGCACGT	20.983/21.049	20.748/20.717

* Ct values are presented in duplicates for HIV-free and HIV + Mengovirus samples.

Supplementary Table S8: Performance Metrics of 12 qPCR Primers for Mengovirus Analysis.

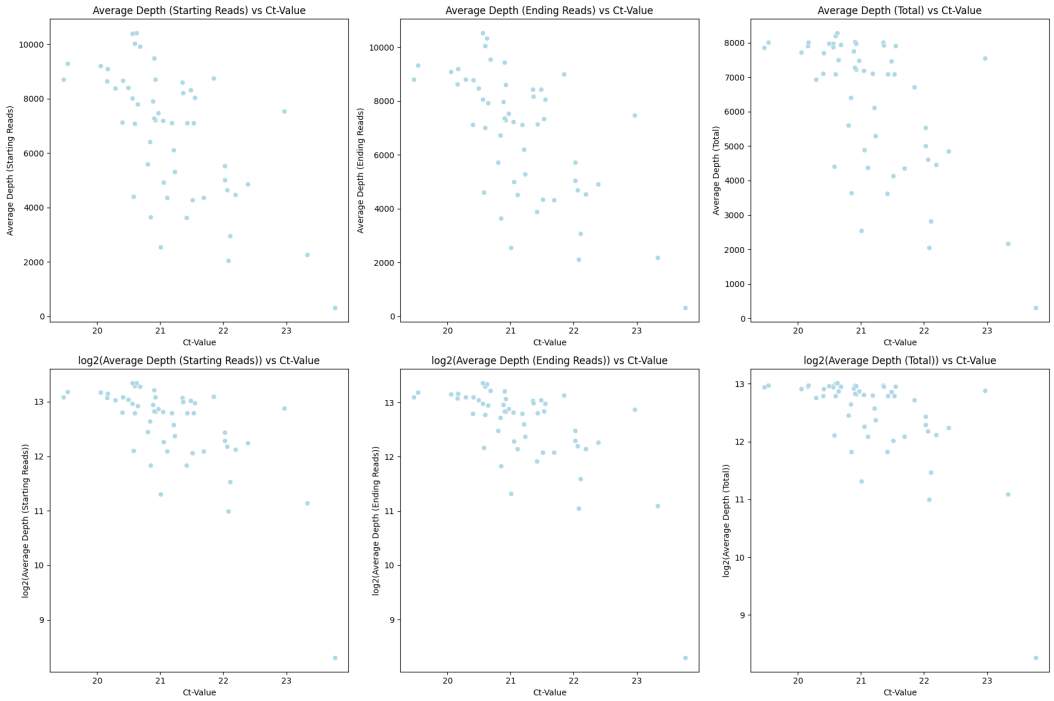
Primer Pair	Slope PX	Intercept PX	R^2 PX	Std Err PX	Efficiency (%) PX
07	-1.037	8.640	0.997	0.023	95.069
11	-1.073	9.415	0.994	0.039	90.766
15	-1.022	9.822	0.996	0.030	97.068
17	-1.037	8.655	0.991	0.043	95.168
21	-1.011	9.319	0.995	0.033	98.437
24	-1.023	8.344	0.992	0.040	96.871
27	-1.038	8.422	0.989	0.048	94.991
33	-1.025	8.542	0.992	0.042	96.609
37	-1.037	8.878	0.992	0.042	95.079
40	-1.026	8.105	0.995	0.033	96.462
56	-1.095	10.816	0.992	0.043	88.273
Previous	-1.041	9.119	1.000	0.008	94.567

Legend: *Slope PX*: the slope of the standard curve for PX samples; *Intercept PX*: intercept of the standard curve for PX samples; *R^2 PX*: coefficient of determination for PX samples; *Std Err PX*: standard error of the slope for PX samples; *Efficiency (%) PX*: PCR efficiency percentage for PX samples.

Supplementary Table S9: Performance Metrics of 12 qPCR Primers for Mengovirus Analysis.

Primer Pair	**07**	**11**	**15**	**17**	**21**	**24**	**27**	**33**	**37**	**40**	**56**	**Previous Mengovirus Primers**
PX Ranking *	8th	11th	2nd	6th	1st	3rd	9th	4th	7th	5th	12th	10th

* PX samples were extracted with the MagMAX™ Viral/Pathogen (MVP) Nucleic Acid Isolation Kit (Thermo Fisher Scientific).



Supplementary Figure S7: Depth vs Ct-Value (Raw and log-transformed). Top row: Scatter plots depict the relationship between Ct-values and average sequencing depth for starting, ending, and total reads, respectively. The data points exhibit a distribution that suggests a potential negative correlation between Ct-values and sequencing depth. **Bottom Row:** Log2-transformed scatter plots for the same relationships as the top row, revealing more detail in the variance at lower Ct-values and higher sequencing depths.

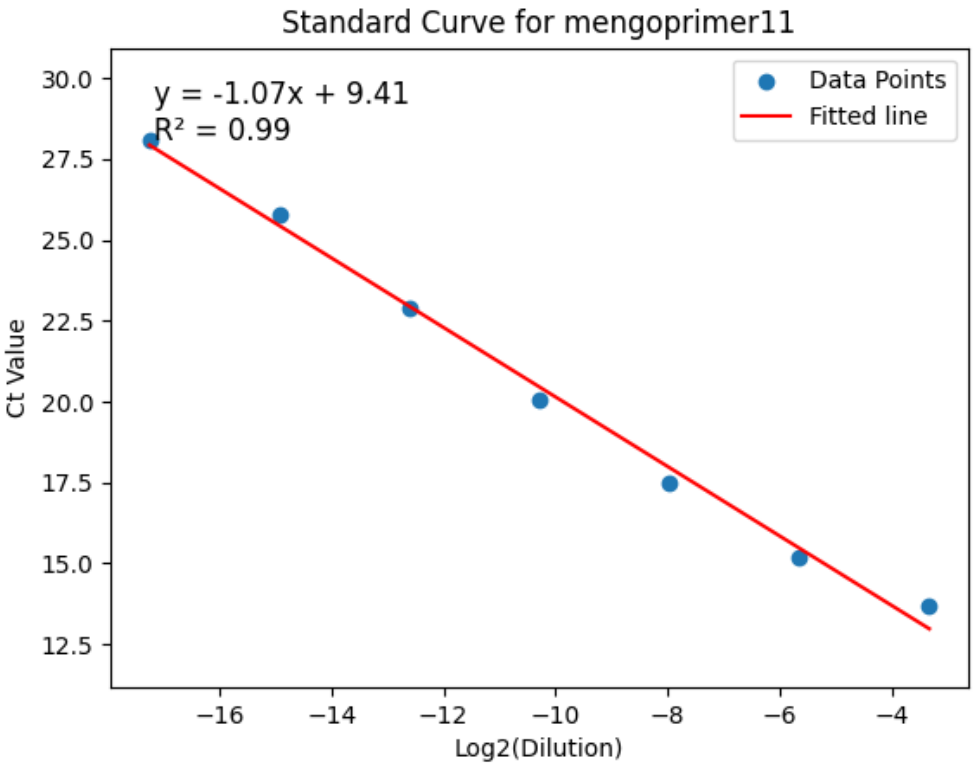
Supplementary Table S10: Summary of qPCR Efficiency and R² Values for PX Samples.

	PX Samples *
Efficiency Standard Deviation	2.82
Median Efficiency (%)	95.12
Mean Efficiency (%)	94.95
Median R²	0.993
Mean R²	0.9938

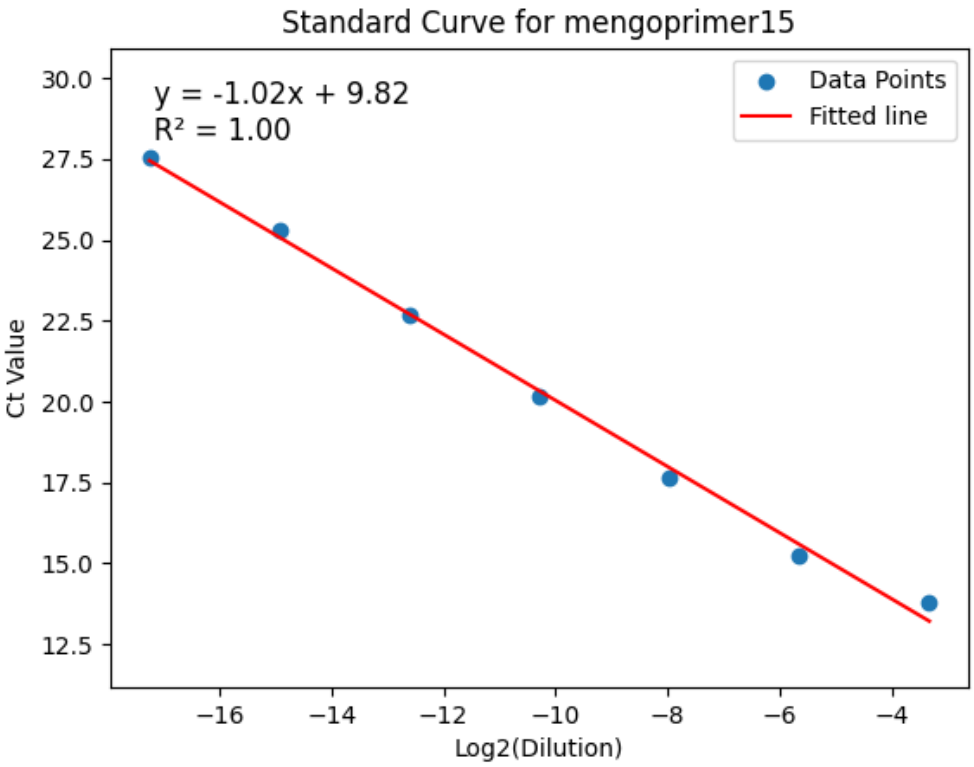
* PX samples were extracted with the MagMAX™ Viral/Pathogen (MVP) Nucleic Acid Isolation Kit (Thermo Fisher Scientific).

Supplementary Table S11: Selected Primers Based on Standard Curve Efficiency.

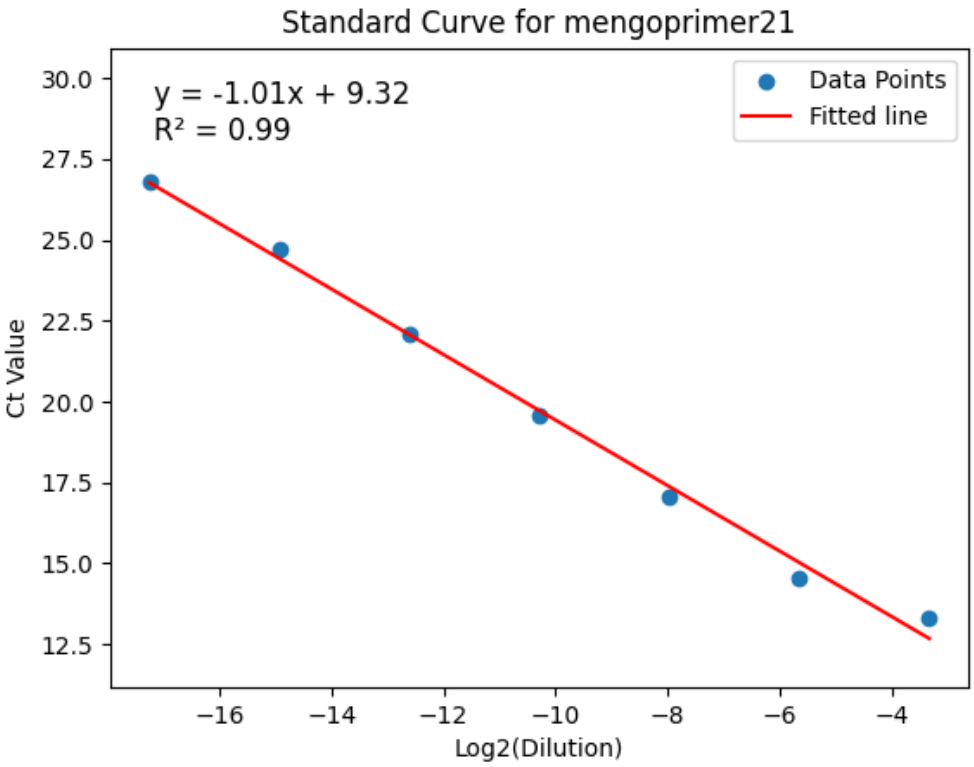
Primer Pair	**Target Region**	**Forward Primer**	**Reverse Primer**
Primer Pair 11	P1 (Protein 1B)	TATCAGGTGAGGATGGTGGTGT	GCGTTCTTGTTCCATTAGGCAG
Primer Pair 15	P1 (Protein 1B)	ACTACGGCATGAGGTACTGTCT	GGGCACTTTGTTCCCAATGAAG
Primer Pair 21	P1 (Protein 1C)	CCCTCAAGGCGTAGAAAATGCA	GGACACGCCTTGTTGGAGAAA
Previous *	P3 (Protein 3C)	GCCTGTTTATTTGCCTGAGAAC	CGACTCTAAGCTCCCAGATTTC



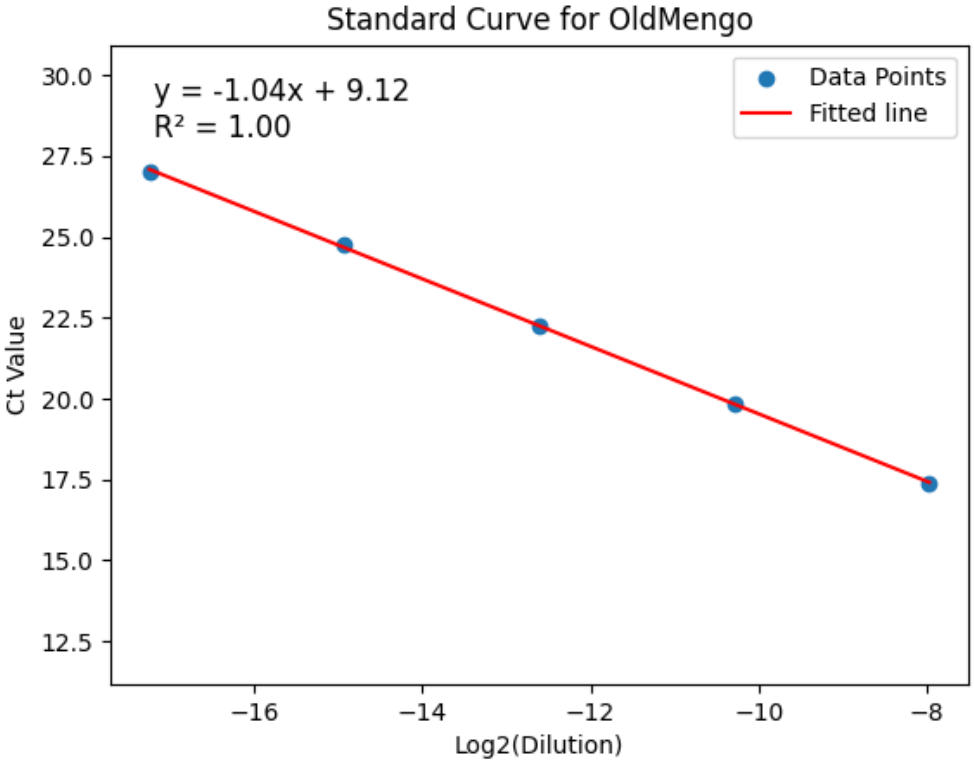
Supplementary Figure S8: Standard Curve for Mengoprimer11. A standard curve displaying Ct-values against log2-transformed dilution factors. The fitted linear regression line (y = -1.07x + 9.41) exhibits a high R^2 value of 0.99, indicating a strong linear relationship between the Ct-values and the dilution factors.



Supplementary Figure S9: Standard Curve for Mengoprimer15. This standard curve shows the Ct-values as a function of log2-transformed dilution factors. The linear regression line (y = -1.02x + 9.82) with an R^2 of 1.00 suggests a highly consistent and predictable performance across the dilution series.



Supplementary Figure S10: Standard Curve for Mengoprimer21. This standard curve plots Ct-values against log2-transformed dilution factors. The linear fit (y = -1.01x + 9.32) and an R^2 value of 0.99 demonstrate a strong correlation and high precision in the assay's quantitative performance.



Supplementary Figure S11: Standard Curve for Previous Mengovirus Primer Pair. The standard curve for shows a linear relationship between Ct-values and log2-transformed dilution factors, with the equation y = -1.04x + 9.12 and an R^2 of 1.00. The consistency in the data indicates reliable quantification across the dilution series.

Supplementary Table S12: R^2 Values and Variance for Ct and ΔCt Datasets.

Ranking	Primer Pair	Dataset	R^2	Error Variance
5th	11	HIV RNA TaqMan HIV Mean Ct (#11)	0.6612032	0.3341402
8th		Library ΔCt HIV TaqMan x Mengovirus SYBR (#11)	0.6413772	0.3646268
12th		HIV Library TaqMan HIV Mean Ct (#11)	0.5964365	0.441237
2nd	**15**	**Library ΔCt HIV TaqMan x Mengovirus SYBR (#15)**	0.670697	0.3201795
3rd		HIV RNA TaqMan HIV Mean Ct (#15)	0.6612032	0.3341402
9th		HIV Library TaqMan HIV Mean Ct (#15)	0.6123064	0.4128993
4th	21	HIV RNA TaqMan HIV Mean Ct (#21)	0.6631901	0.3311854
6th		Library ΔCt HIV TaqMan x Mengovirus SYBR (#21)	0.6423362	0.3631089
10th		HIV Library TaqMan HIV Mean Ct (#21)	0.6059003	0.4241595
1st	Previous	HIV RNA TaqMan HIV Mean Ct (Previous)	0.677653	0.3101991
7th		Library ΔCt HIV TaqMan x Mengovirus SYBR (Previous)	0.6413772	0.3646268
11th		HIV Library TaqMan HIV Mean Ct (Previous)	0.5964365	0.441237

Supplementary Table S13: Regression Coefficients Before and After the Addition of ΔCt15 as a Predictor.

Model	Predictor	Coefficient	Std. Error	t-value	p-value
Simple Linear Regression	Intercept	1.234	0.567	2.17	0.036
	Log (paired-end reads)	0.567	0.123	4.61	<0.001
Multiple Regression	Intercept	0.987	0.456	2.16	0.035
	Log (paired-end reads)	0.456	0.098	4.65	<0.001
	Library ΔCt HIV TaqMan x Mengovirus SYBR (#15)	0.234	0.078	3	0.004

Supplementary Table S14: Confidence Intervals of Multiple Regression Predictors.

Predictor	2.50%	97.50%
(Intercept)	3.493773	5.36069711
Library ΔCt HIV TaqMan x Mengovirus SYBR (#15)	-0.23055	-0.09907608
Log (paired-end reads)	0.136313	0.48131122

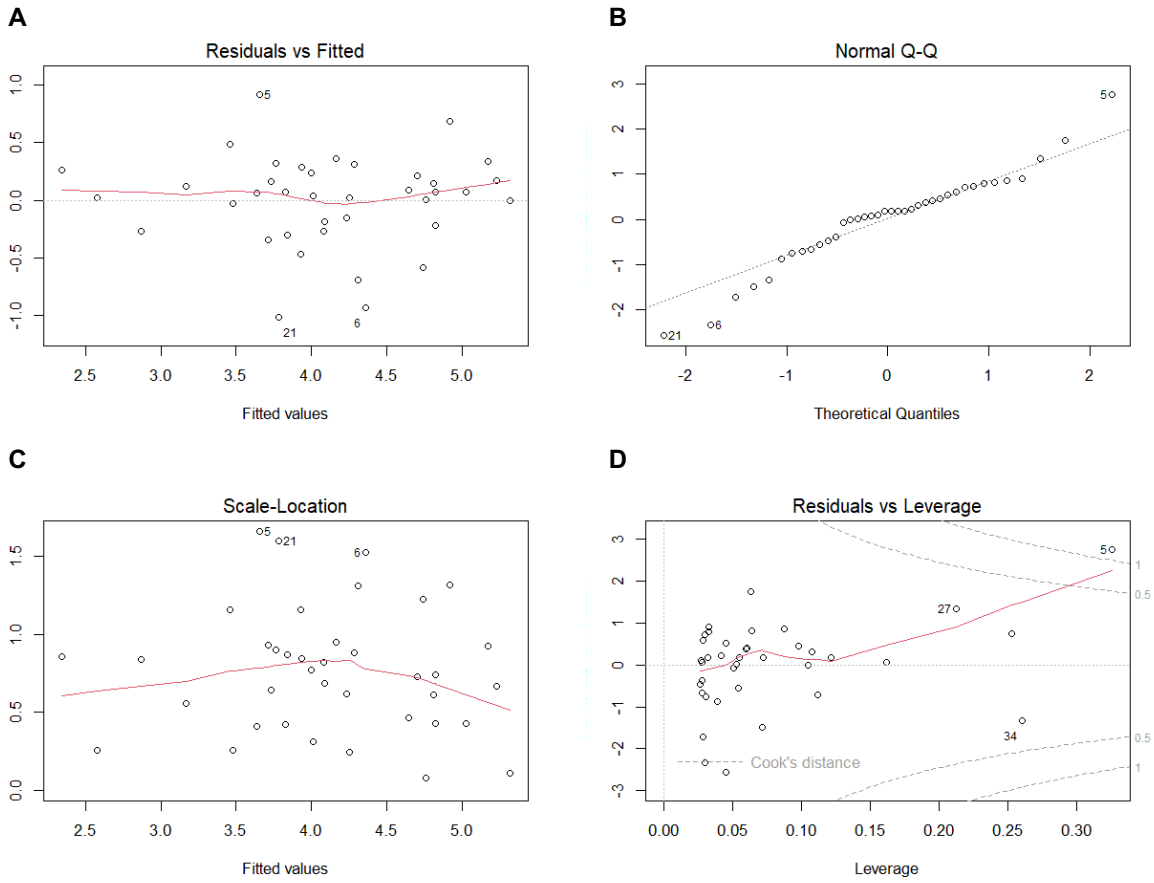


Figure S12: Diagnostic Plots for the Multiple Linear Regression Model of Two Predictors. (A) Residuals vs Fitted: Residuals are plotted against fitted values. The red LOESS line highlights any patterns in the residuals. Random scatter around the horizontal line indicates a well-fitting model. **(B) Normal Q-Q:** The quantiles of the residuals are compared to a normal distribution. Points along the dashed line suggest the normality of the residuals. **(C) Scale-Location (Spread-Location):** The square root of the absolute standardised residuals is plotted against the fitted values. The red LOESS line indicates the trend. A horizontal trend suggests homoscedasticity, while a pattern suggests heteroscedasticity. **(D) Residuals vs Leverage:** Residuals are plotted against leverage to identify influential data points. The red LOESS line shows the trend. Dashed lines represent Cook's distance, with points outside these lines indicating highly influential observations.