Supplementary Information

Enzymatic Colorimetric Encoding-Based Digital Medicine for Pancreatic Cancer Diagnosis

Dongsheng Mao[1,#], Chenbin Liu[1,2,#], Runchi Zhang[1,2,#], Zhiyuan Ma[3], Liang Wu[1], Mingjin Zhu[1,2], Xiaochen Tang[4], Wen Chen[1,2], Jie Deng[1,2], Hongquan Gou[1,2], Xiong Dun[3], Jingqi Chen[5], Zhaocheng Liu[6], Wenxing Li[1,*], Fenyong Sun[1,*], Xiaoli Zhu[1,*]

1 Department of Clinical Laboratory Medicine, Shanghai Tenth People's Hospital, School of Medicine, Tongji University, Shanghai 200072, China.
2 School of Medicine, Tongji University, Shanghai 200331, China.
3 Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China.
4 Department of Clinical Laboratory Medicine, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China.
5 Department of Clinical Laboratory Medicine, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai 200434, China.
6 Department of Laboratory Medicine, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi People's Hospital, Wuxi Medical Center, Nanjing Medical University, Wuxi 214023, China.
# These authors contributed equally to this work.

Table of Contents

Supplementary Figures

Supplementary Tables

Supplementary Table 1. Inter-rater agreement in subjective visual interpretation of EnCODE colorimetric readouts.

Supplementary Table 2. Target detection performance of three weighting schemes at a fixed weighting coefficient.

Supplementary Table 3. Weighting coefficient responsiveness of the three weighting schemes at a fixed target concentration.

Supplementary Table 4. Summary of baseline clinical characteristics of the training cohort.

Supplementary Table 5. Summary of baseline clinical characteristics of the validation cohort.

Supplementary Table 6. DNA/RNA sequences and modifications for dual miRNA detection based on EnCODE.

Supplementary Table 7. DNA/RNA sequences and modifications for EnCODE performance validation.

Supplementary Table 8. DNA/RNA sequences and modifications for fluorescence-based EnCODE performance validation.

Supplementary Table 9. DNA/RNA sequences and modifications for multi-miRNA detection based on EnCODE.

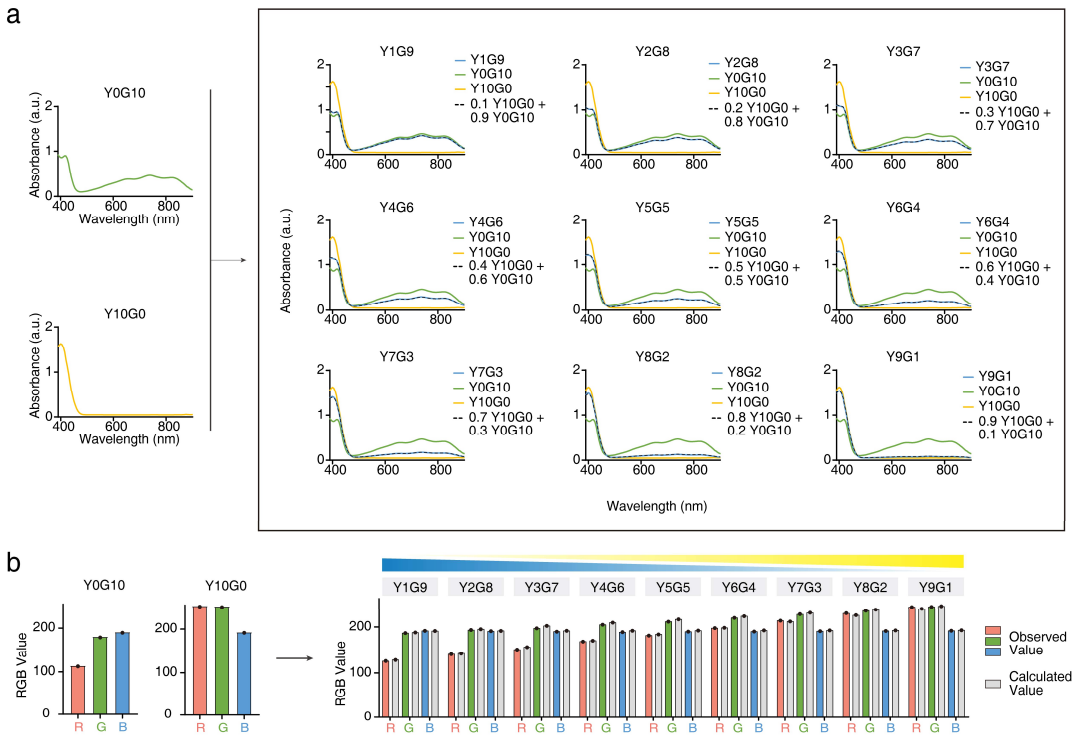
Supplementary Table 10. DNA/RNA sequences and modifications for weighted calculation based on EnCODE.

Supplementary Table 11. DNA sequences and modifications for clinical sample detection based on EnCODE.

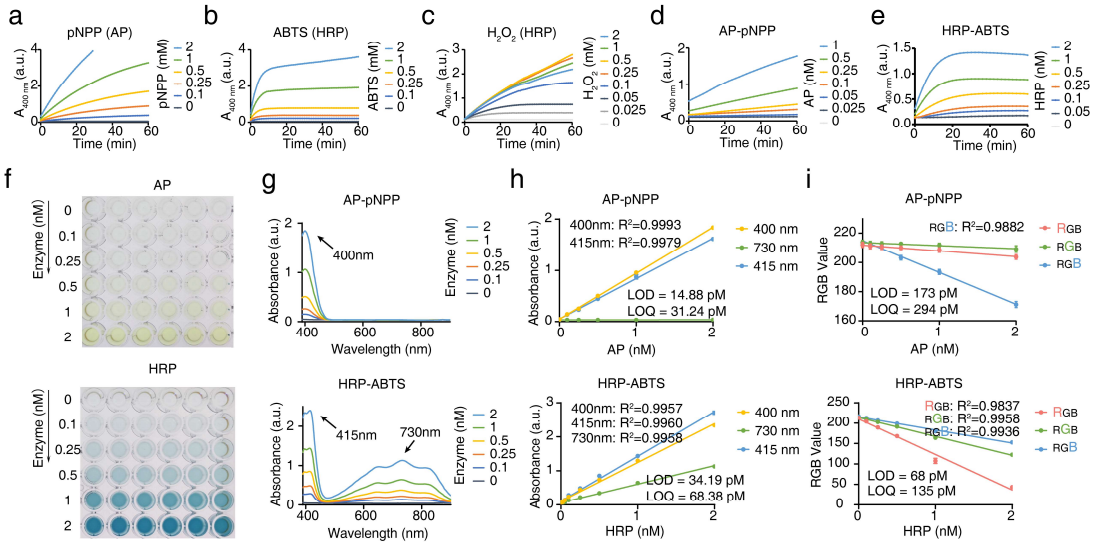
Computer Code

Uncropped scans of all gels in Supplementary Figures

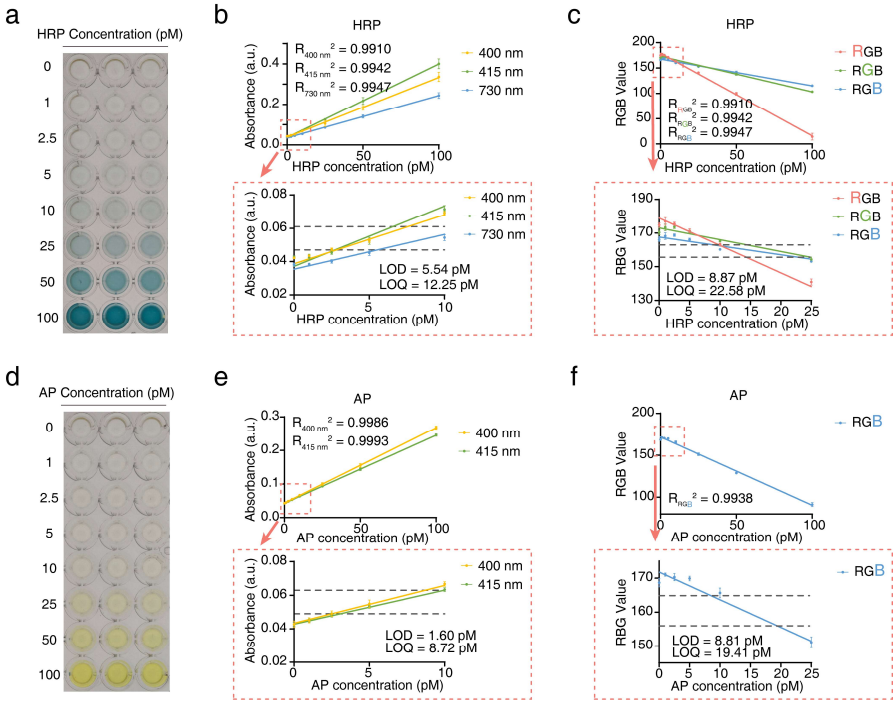
Supplementary Figures



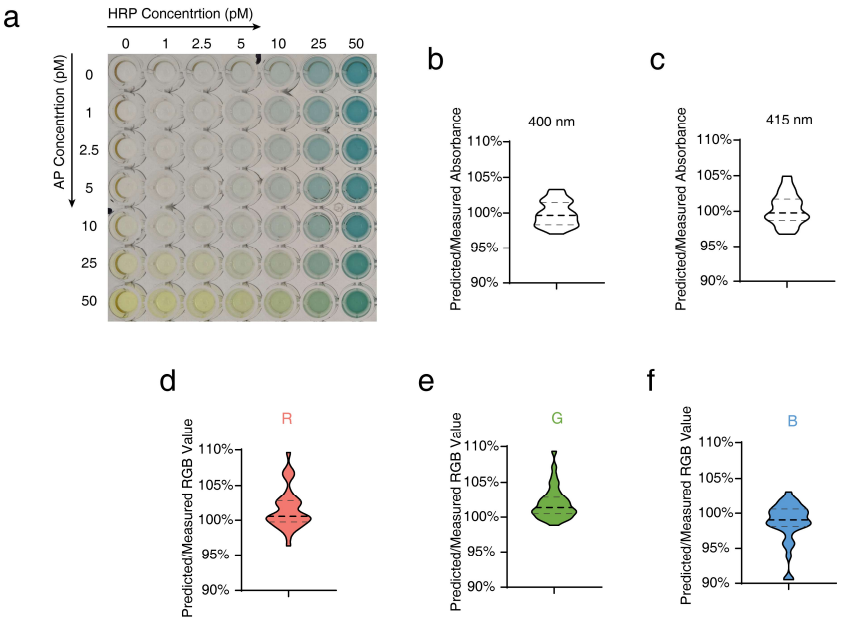
Supplementary Fig. 1. Absorption spectra and RGB values of dual-color proportion encoded solutions. (a) Graphs illustrating the relationship between the absorption spectra of yellow and green original solutions and those of encoded solutions in different proportions. (b) Graph illustrating the relationship between the RGB values of yellow and green original solutions and those of encoded solutions in different proportions. Source data are provided as a Source Data file.



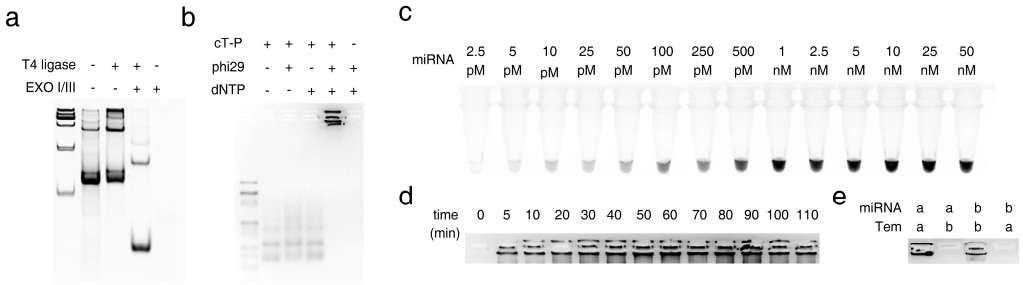
Supplementary Fig. 2. Enzyme-catalyzed colorimetric reaction performance validation. (a) Enzyme kinetics verification with AP catalyzed under different pNPP concentrations. (b) Enzyme kinetics verification with HRP catalyzed under different ABTS concentrations. (c) Enzyme kinetics verification with HRP catalyzed under different H_2O_2 concentrations. (d) Enzyme kinetics verification with different AP concentrations under optimal reaction conditions. (e) Enzyme kinetics verification with different HRP concentrations under optimal reaction conditions. (f) Photograph of the enzyme-catalyzed colorimetric reaction with different concentrations of AP and HRP. (g) Absorption spectra of the enzyme-catalyzed colorimetric reaction with different concentrations of AP and HRP. (h) Plot of the absorbance values of the characteristic peaks in the absorption spectra after the enzyme-catalyzed colorimetric reaction versus the enzyme concentration. (i) Plot of the RGB values after the enzyme-catalyzed colorimetric reaction versus the enzyme concentration. Data are presented as the mean ± s.d. from six independent experiments ($n = 6$). Source data are provided as a Source Data file.



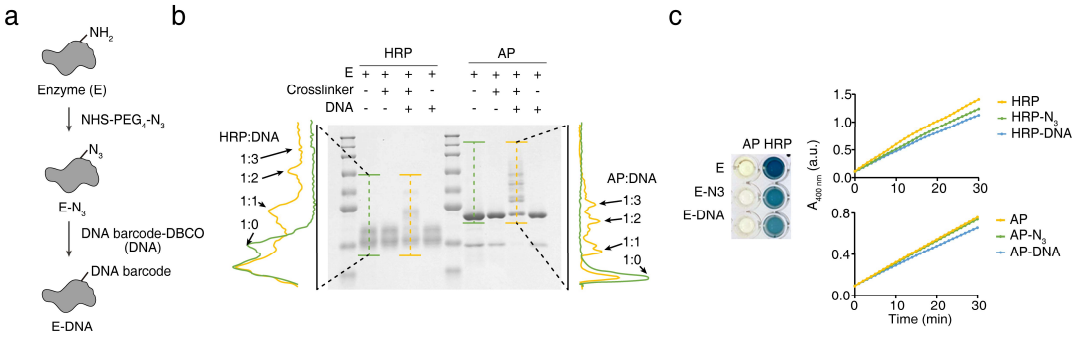
Supplementary Fig. 3. Performance of enzyme-catalyzed colorimetric reactions at ultralow enzyme concentrations with a 1 h development time. (a) Photographs of the colorimetric reaction at different HRP concentrations. (b) Relationship between the characteristic absorbance peak intensity in the UV–vis spectra and HRP concentration. (c) Relationship between RGB values after color development and HRP concentration. (d) Photographs of the colorimetric reaction at different AP concentrations. (e) Relationship between the characteristic absorbance peak intensity in the UV–vis spectra and AP concentration. (f) Relationship between RGB values after color development and AP concentration. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Source data are provided as a Source Data file.



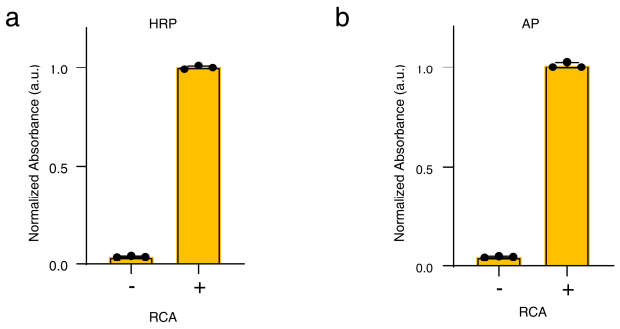
Supplementary Fig. 4. Validation of color and spectral additivity in dual-enzyme catalyzed colorimetric reactions at ultralow enzyme concentrations. (a) Photographs of the 1 h colorimetric reactions catalyzed by mixtures of HRP and AP at different concentration combinations. (b) Ratio of predicted-to-measured absorbance at 400 nm for each well in panel (a). *n*=49. (c) Ratio of predicted-to-measured absorbance at 415 nm for each well in panel (a). *n*=49. (d) Ratio of predicted-to-measured R-channel values for each well in panel (a). *n*=49. (e) Ratio of predicted-to-measured G-channel values for each well in panel (a). *n*=49. (f) Ratio of predicted-to-measured B-channel values for each well in panel (a). Predicted-to-measured ratios, where predicted values were calculated as the linear sum of the corresponding single-enzyme responses. *n*=49. Source data are provided as a Source Data file.



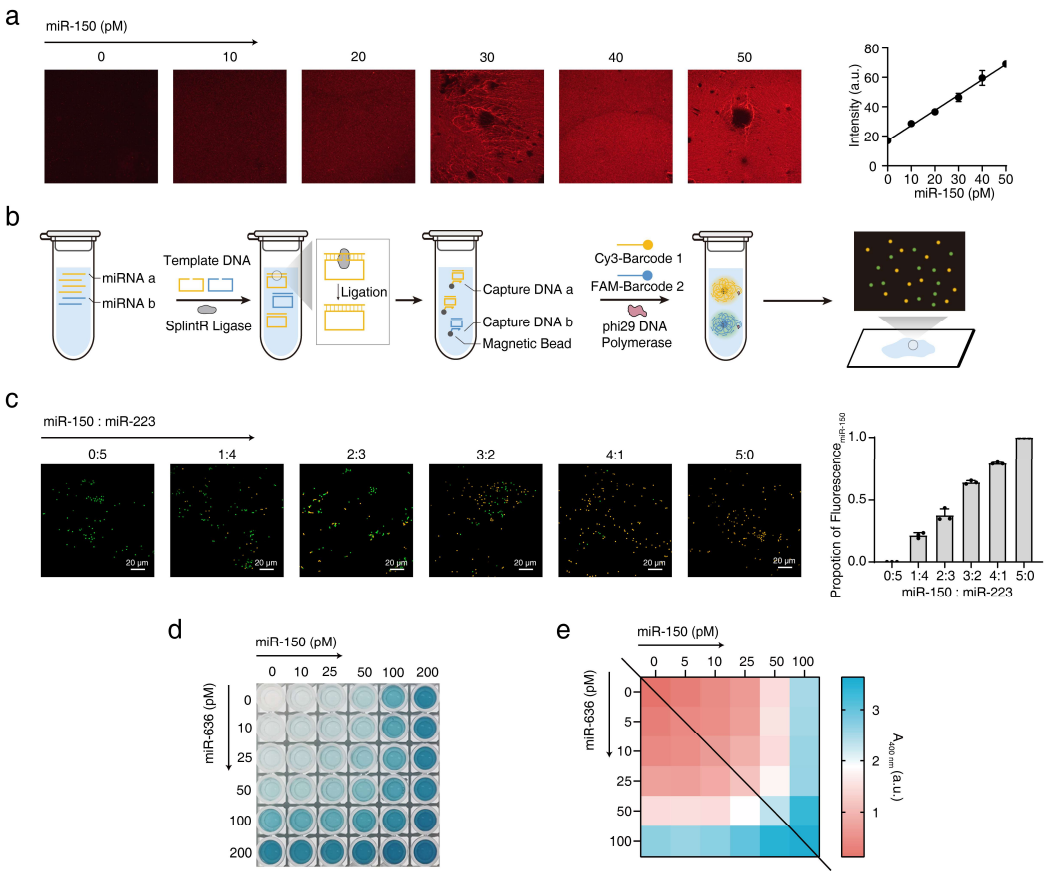
Supplementary Fig. 5. Validation of miRNA rolling circle amplification. (a) Agarose gel electrophoresis image of the successful circularization of the template DNA using miRNA as primers. (b) Agarose gel electrophoresis image of the successful rolling circle amplification achieved with miRNA as primers. (c) Fluorescence photograph of rolling circle amplification performed with miRNA primers at different concentrations. (d) Agarose gel electrophoresis image of rolling circle amplification conducted for different durations. (e) Agarose gel electrophoresis image of rolling circle amplification performed with different combinations of two pairs of miRNA primers and template DNA. The experiments were independently repeated three times with similar results. Uncropped scans of gels are supplied at the end of the Supplementary Information file.



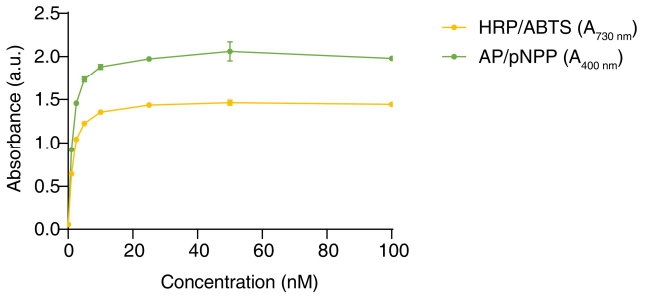
Supplementary Fig. 6. Verification of the cross-linking of enzymes with DNA barcodes. (a) Schematic diagram of the cross-linking process between enzymes and DNA barcodes. (b) SDS-PAGE electrophoresis image of the successful cross-linking between AP and HRP with DNA barcodes. (c) Enzymatic kinetics analysis of the enzyme before and after its cross-linking with DNA barcodes. The experiments were independently repeated three times with similar results. Uncropped scans of gels are supplied at the end of the Supplementary Information file. Source data are provided as a Source Data file.



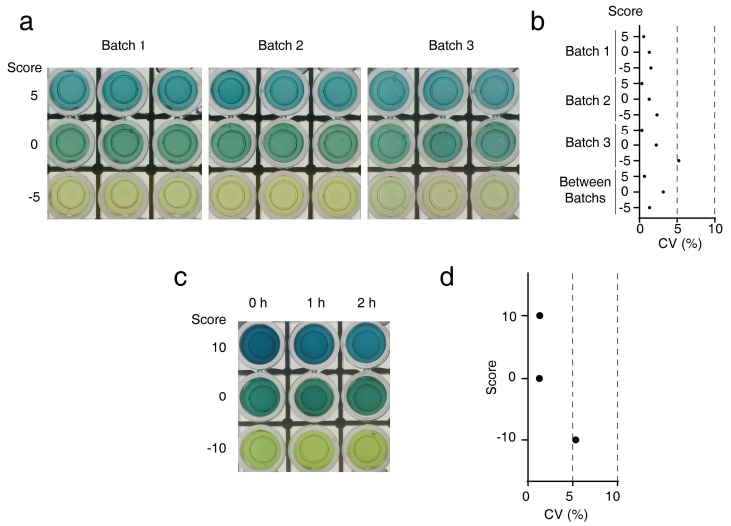
Supplementary Fig. 7. Estimation of enzyme loading per RCA product by comparing RCA-enabled capture versus direct capture. (a) Comparison of normalized absorbance obtained by directly capturing the HRP–DNA conjugate on the capture-strand–coated plate (no RCA) versus capturing the same conjugate via RCA products after a 30-min RCA step (RCA). (b) Comparison of normalized absorbance obtained by directly capturing the AP–DNA conjugate on the capture-strand–coated plate (no RCA) versus capturing the same conjugate via RCA products after a 30-min RCA step (RCA). Data are presented as the mean ± s.d. from three independent experiments (n = 3). Source data are provided as a Source Data file.



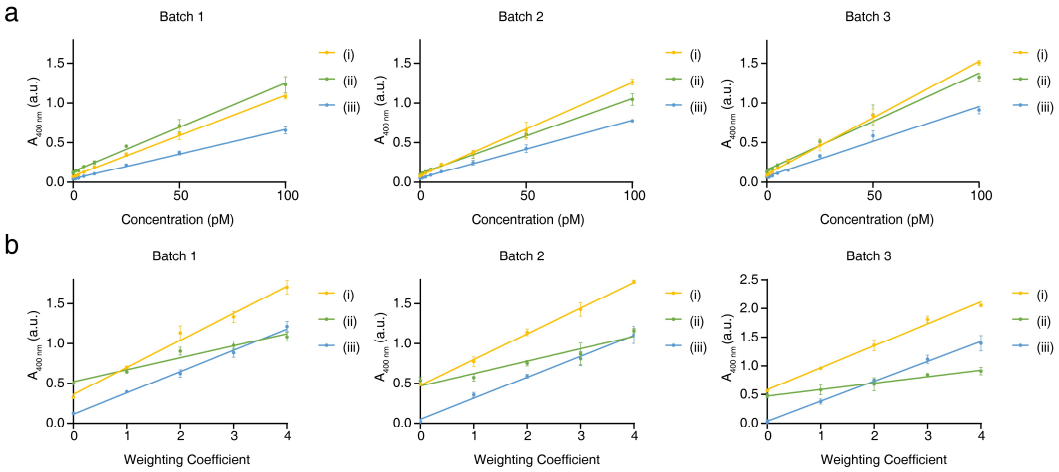
Supplementary Fig. 8. Validation of the multiplexed detection of miRNAs using the EnCODE. (a) Verification of the feasibility of enzymatic colorimetric encoding detection by replacing the colorimetric enzyme-DNA barcodes with fluorescently labeled DNA barcodes. Display the fluorescence results for different target concentrations, as well as the plot of target concentration versus fluorescence intensity. (b) Schematic representation of the feasibility validation of the enzymatic colorimetric encoding system for multiplexed detection on magnetic beads. (c) Demonstration of multiplexed detection on magnetic beads using the enzymatic colorimetric encoding system, showing the fluorescence results for different proportions of targets, and the plot of the proportion of different fluorescent magnetic beads versus the proportion of different targets. (d) In the enzymatic colorimetric detection system, the results of detecting the mimics of miR-150 and miR-636 were both indicated by green color. The results are presented with photographs. (e) The heat map of absorbance for panel (d). The experiments were independently repeated three times with similar results. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Source data are provided as a Source Data file.



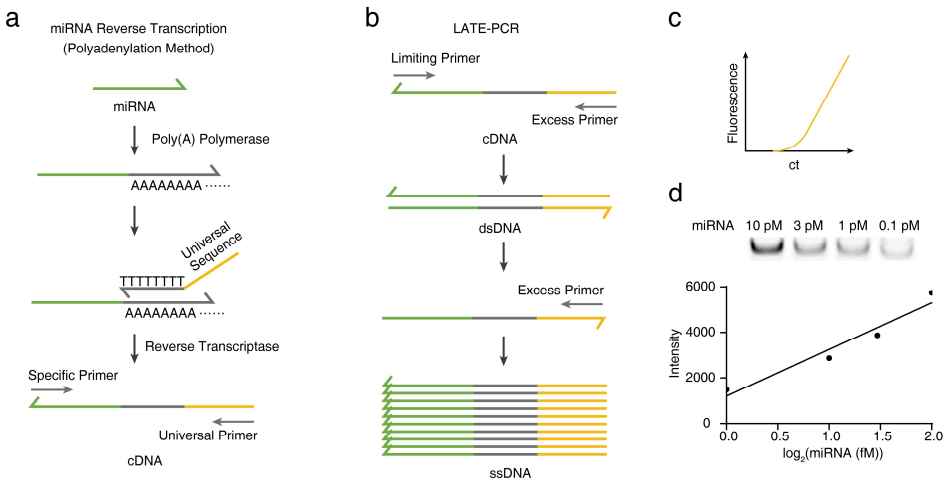
Supplementary Fig. 9. Effect of enzyme–DNA conjugate concentration on EnCODE readout at a fixed target level. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Source data are provided as a Source Data file.



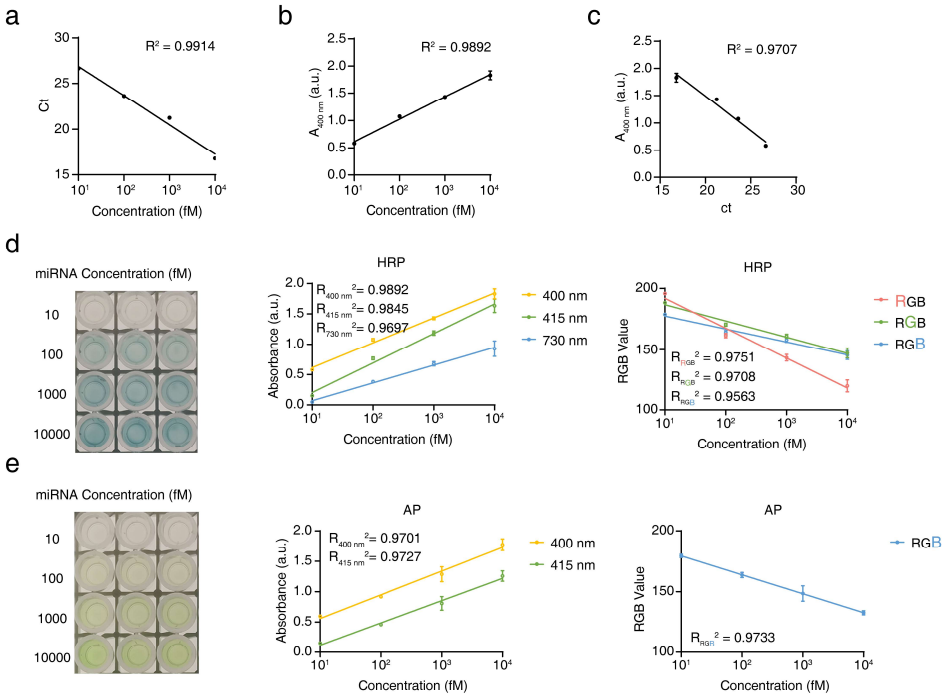
Supplementary Fig. 10. Stability assessment of the colorimetric-to-risk-score transformation. (a) Representative photographs of EnCODE readouts for samples spanning different risk-score levels measured by different operators across independent batches. (b) Quantification of well-to-well and batch-to-batch variability of the chromatic parameters in panel (a). (c) Representative photographs of EnCODE readouts for samples at different risk-score levels following defined room-temperature holding times prior to processing. (d) Coefficient of variation (CV%) analysis of the readouts in panel (c) across different holding times and risk-score levels. Source data are provided as a Source Data file.



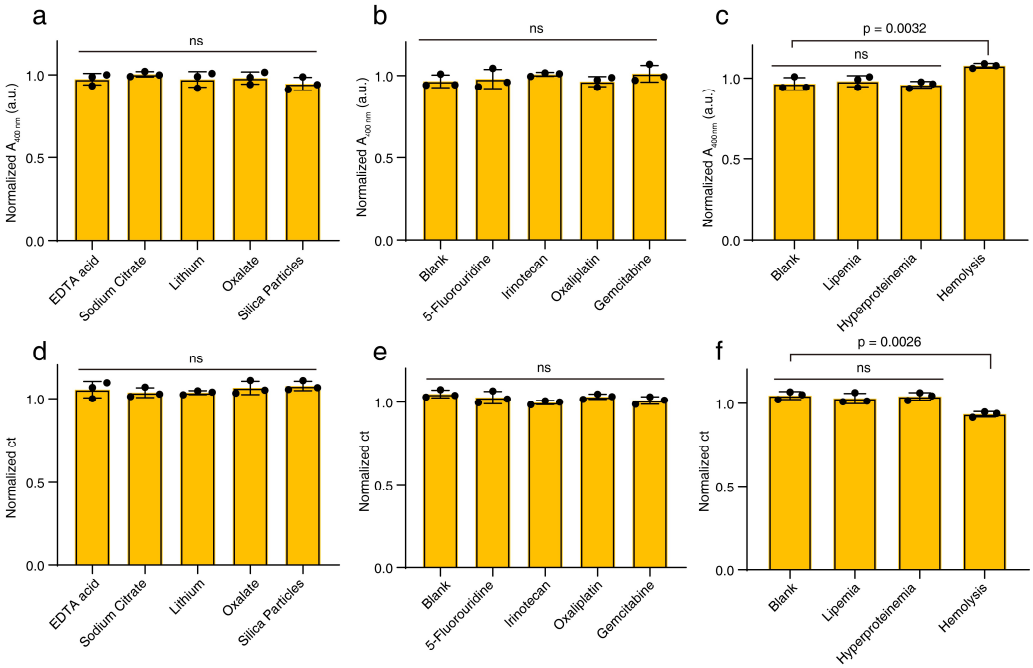
Supplementary Fig. 11. Multi-batch performance evaluation of three weighting schemes. (a) The relationship between target concentration and absorbance response at a fixed weighting coefficient for the three weighting schemes. (b) Weighting responsiveness evaluated at a fixed target concentration, showing the assay output as a function of the weighting coefficient for the three weighting schemes. Data are presented as mean ± s.d. from three independent batches (n = 3). Source data are provided as a Source Data file.



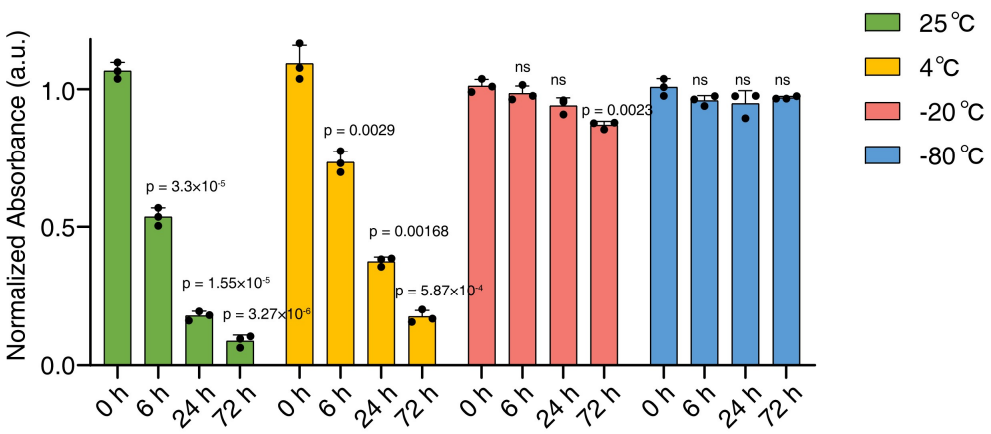
Supplementary Fig. 12. Validation of the Pre-amplification Performance of miRNAs. (a) Schematic representation of the reverse transcription of miRNAs using a tailing method. (b) Schematic representation of the LATE-PCR performed after the reverse transcription of miRNAs. (c) Schematic representation of the relationship between the amplified products in the LATE-PCR and the ct values. (d) PAGE electrophoresis image of the LATE-PCR products from different concentrations of miRNAs, with the plot of the gray values versus the logarithm of the miRNA concentrations. The experiments were independently repeated three times with similar results. Uncropped scans of gels are supplied at the end of the Supplementary Information file. Source data are provided as a Source Data file.



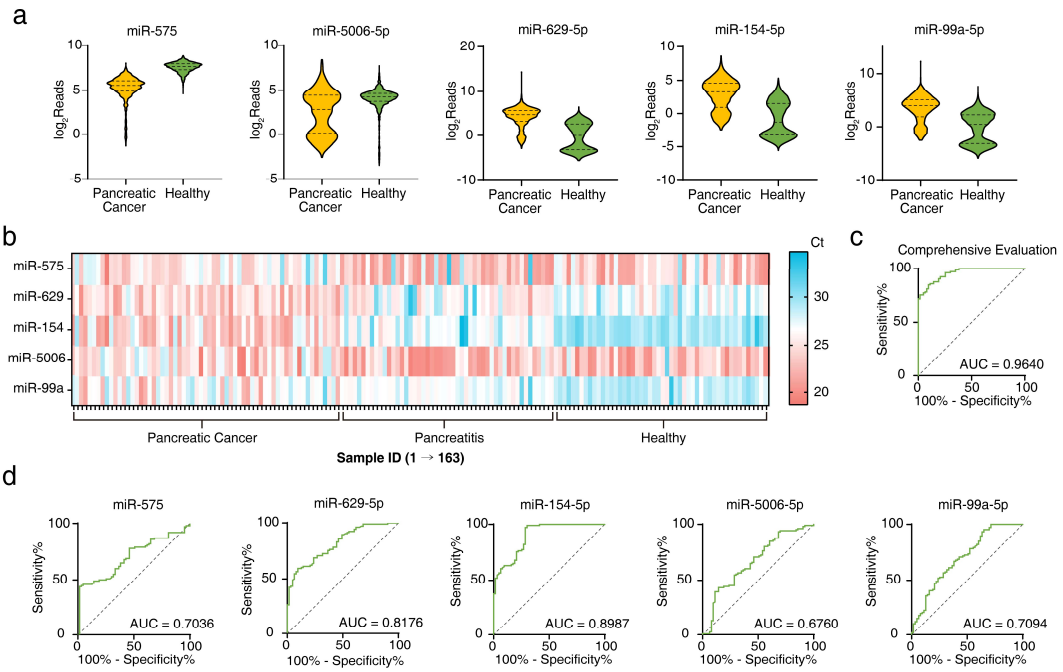
Supplementary Fig. 13. Analytical benchmarking of EnCODE for miRNA quantification. (a) Correlation between input miRNA concentrations (processed with the same pre-amplification step as in the workflow) and qRT-PCR quantification results. (b) Correlation between input miRNA concentrations and EnCODE readouts for the same sample set. (c) Comparison of miRNA quantification results obtained by qRT-PCR and EnCODE for matched samples. (d) Performance validation of EnCODE for miRNA detection using the HRP–ABTS system. Left: representative photographs of EnCODE colorimetric outputs at different miRNA concentrations; middle: relationship between absorbance at characteristic spectral peaks and miRNA concentration; right: relationship between RGB values and miRNA concentration after color development. (e) Performance validation of EnCODE for miRNA detection using the AP–pNPP system. Left: representative photographs of EnCODE colorimetric outputs at different miRNA concentrations; middle: relationship between absorbance at characteristic spectral peaks and miRNA concentration; right: relationship between RGB values and miRNA concentration after color development. Data are presented as mean ± s.d. from three independent experiments (n = 3). Source data are provided as a Source Data file.



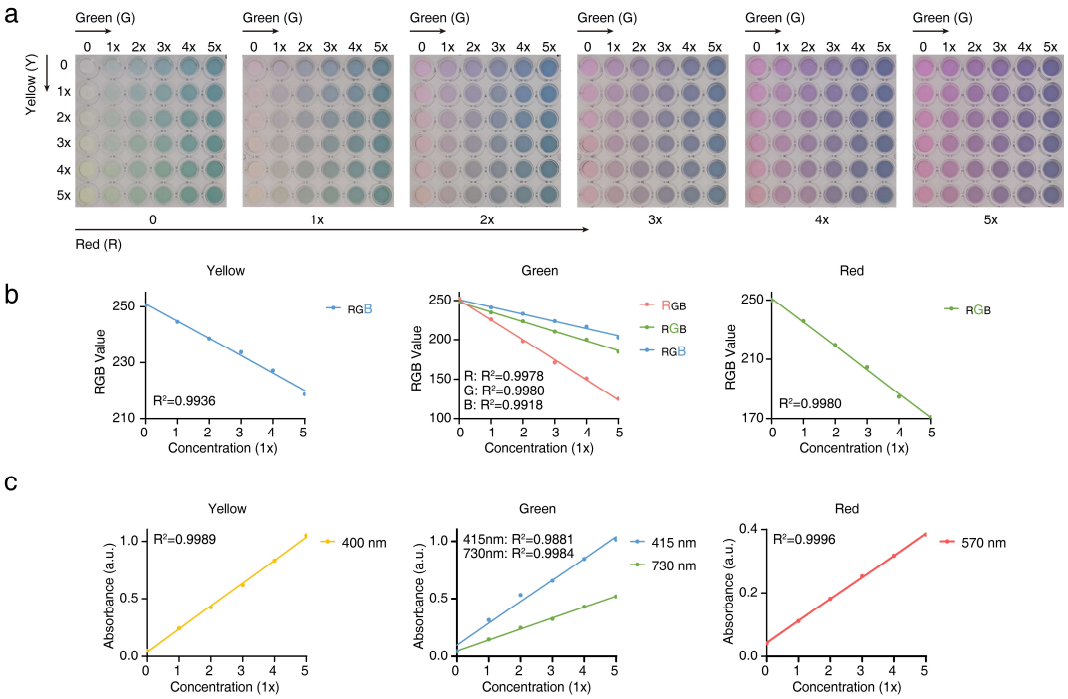
Supplementary Fig. 14. Interference validation of the EnCODE platform. (a) EnCODE miRNA readouts for blood aliquots collected under different anticoagulant/coagulant conditions, reported as normalized absorbance. (b) EnCODE miRNA readouts (normalized absorbance) for matched blood aliquots with or without representative pancreatic cancer treatment drugs. (c) EnCODE miRNA readouts (normalized absorbance) for matched blood aliquots under simulated matrix interferences, including hemolysis, lipemia, and elevated protein content, generated by adding equal volumes of red-blood-cell lysate, Intralipid, or albumin solution, respectively, alongside an untreated control. (d–f) Corresponding qRT-PCR results for the same sample sets shown in (a–c), reported as normalized Ct values. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Statistical analysis was performed using a Two-tailed Student's t-test, and p values are indicated. P values were adjusted for multiple comparisons within each figure using the Benjamini–Hochberg procedure (FDR=0.05). Source data are provided as a Source Data file.



Supplementary Fig. 15. Impact of sample storage conditions and duration on miRNA detection results using EnCODE. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Statistical analysis was performed using a Two-tailed Student's t-test, and p values are indicated. P values were adjusted for multiple comparisons within each figure using the Benjamini–Hochberg procedure (FDR=0.05). Source data are provided as a Source Data file.



Supplementary Fig. 16. Diagnostic model and PCR-based performance benchmarking. (a) Differential expression analysis of the five selected miRNA markers in the GSE211692 serum dataset. (b) miRNA expression profiles measured in the expanded clinical cohort, including 63 pancreatic cancer serum samples, 50 pancreatitis serum samples, and 50 healthy serum samples. (c) ROC curve showing the classification performance of the conventional PCR-based workflow for pancreatic cancer detection using the five-miRNA model. (d) ROC curves showing the classification performance of PCR-based single-target miRNA assays for pancreatic cancer detection. Source data are provided as a Source Data file.



Supplementary Fig. 17. Three-dimensional colorimetric performance verification. (a) The original photographs of the colorimetric encoding results obtained after mixing different series concentrations of the green, yellow, and red stock solutions. (b) The plot of the concentrations of single-color solutions versus the absorbance values of their characteristic peaks in the absorption spectra. (c) The plot of the concentrations of single-color solutions versus the RGB values of their colors. Source data are provided as a Source Data file.

Supplementary Tables

Supplementary Table 1. Inter-rater agreement in subjective visual interpretation of EnCODE colorimetric readouts.

Sample (reference category)	Rater A	Rater B	Rater C
1 (Strong positive)	Positive	Positive	Positive
2 (Strong positive)	Positive	Positive	Positive
3 (Strong positive)	Positive	Positive	Positive
4 (Weak positive)	Positive	Positive	Positive
5 (Weak positive))	Positive	Positive	Positive
6 (Weak positive))	Positive	Positive	Positive
7 (Weak negative)	Negative	Positive	Negative
8 (Weak negative)	Negative	Negative	Positive
9 (Weak negative)	Negative	Negative	Negative
10 (Strong negative)	Negative	Negative	Negative
11 (Strong negative)	Negative	Negative	Negative
12 (Strong negative)	Negative	Negative	Negative

Supplementary Table 2. Target detection performance of three weighting schemes at a fixed weighting coefficient.

Metric	scheme (i): Tem1/Tem0 ratio control	scheme (ii): Barcode copy-number modulation	scheme (iii): Competitive binding scheme
LOD*	1.7201 pM	10.1747 pM	2.0720 pM
Sensitivity (slope)	0.0102	0.0111	0.0062
Linear correlation (R^2)	0.9936	0.9860	0.9921
Intra-batch CV%	＜5%	＜5%	＜5%
Inter-batch CV%	＜15%	＜15%	＜15%

* LOD was calculated based on the 3σ criterion unless otherwise specified.

Supplementary Table 3. Weighting coefficient responsiveness of the three weighting schemes at a fixed target concentration.

Metric	scheme (i): Tem1/Tem0 ratio control	scheme (ii): Barcode copy-number modulation	scheme (iii): Competitive binding scheme
LOD*	0.0653	0.2303	0.2059
Sensitivity (slope)	0.3394	0.1504	0.2668
Linear correlation (R^2)	0.9779	0.9361	0.9873
Intra-batch CV%	＜5%	＜5%	＜5%
Inter-batch CV%	＜15%	＜15%	＜15%

* LOD was calculated based on the 3σ criterion unless otherwise specified.

Supplementary Table 4. Summary of baseline clinical characteristics of the training cohort.

Clinical Characteristics	**Pancreatic Cancer**	**Pancreatitis**	**Healthy**
Case	63	50	50
Age	80-90: 6 70-80: 29 60-70: 20 50-60: 8	80-90: 10 70-80: 15 60-70: 7 50-60: 11 40-50: 4 30-40: 3	70-80: 8 60-70: 14 50-60: 8 40-50: 11 30-40: 9
Sex	Male: 30 Female: 33	Male: 25 Female: 25	Male: 25 Female: 25
Disease stage	pancreatic ductal adenocarcinoma :58 pancreatic neuroendocrine tumors: 3 invasive intraductal papillary mucinous neoplasm: 1 mucinous cystic neoplasm:1	/	/
Pathological type	Stage I: 8 Stage II: 18 Stage IV: 37	/	/
Chemotherapy received	Yes: 39 No: 24	/	/

Supplementary Table 5. Summary of baseline clinical characteristics of the validation cohort.

Clinical Characteristics	**Pancreatic Cancer**	**Pancreatitis**	**Healthy**
Case	20	15	15
Age	80-90: 1 70-80: 7 60-70: 6 50-60: 6	70-80: 3 60-70: 6 50-60: 5 40-50: 1	70-80: 1 60-70: 7 50-60: 4 40-50: 3
Sex	Male: 9 Female: 11	Male: 7 Female: 8	Male: 7 Female:8
Disease stage	pancreatic ductal adenocarcinoma :19 pancreatic neuroendocrine tumors: 1	/	/
Pathological type	Stage I: 1 Stage II: 5 Stage IV: 14	/	/
Chemotherapy received	Yes: 12 No: 8	/	/

Supplementary Table 6. DNA/RNA sequences and modifications for dual miRNA detection based on EnCODE.

Name	Sequences (5'-3')	Modification	Nucleic Acid Types
Biotin-Capture DNA	AATCGTGCTTGCCTGATAACCT	5’-biotin	DNA
Target 1	UUCAAGUAAUUCAGGAUAGGU	\	RNA
Target 2	CAUUAUUACUUUUGGUACGCG	\	RNA
Tem 1	AATTACTTGAAAGTGTACTCGCCAAA AGTGTACTCGCCAAAAAGGTTATCAG GCAAGCACGAAAACCTATCCTG	5’-PO4	DNA
Tem 2	AGTAATAATGAATCAAGTAATCCAGA ATCAAGTAATCCAGAAAGGTTATCAG GCAAGCACGAACGCGTACCAAA	5’-PO4	DNA
Barcode 1 (HRP)	GTGTACTCGCCAA	3’-DBCO	DNA
Barcode 2 (AP)	TCAAGTAATCCAG	3’-DBCO	DNA

Supplementary Table 7. DNA/RNA sequences and modifications for EnCODE performance validation.

Name	Sequences (5'-3')	Modification	Nucleic Acid Types
Biotin-Capture DNA	AATCGTGCTTGCCTGATAACCT	5’-biotin	DNA
miR-150	UCUCCCAACCCUUGUACCAGUG	\	RNA
miR-145	GUCCAGUUUUCCCAGGAAUCCCU	\	RNA
miR-636	UGUGCUUGCUCGUCCCGCCCGCA	\	RNA
miR-223	UGUCAGUUUGUCAAAUACCCCA	\	RNA
Tem 150	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 145	GGAAAACTGGACGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAAAGGGATTCCTG	5’-PO4	DNA
Tem 636	CGAGCAAGCACAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAATGCGGGCGGGA	5’-PO4	DNA
Tem 223	ACAAACTGACAAGTGTACTCGCCAAA AGTGTACTCGCCAAAAAGGTTATCAG GCAAGCACGAATGGGGTATTTG	5’-PO4	DNA
Barcode HRP	GTGTACTCGCCAA	3’-DBCO	DNA
miR-150 (single base mismatch)	UCUCCCAACCGUUGUACCAGUG	\	RNA
miR-150 (double base mismatch)	UCUCCCAACCGAUGUACCAGUG	\	RNA
miR-150 (triple base mismatch)	UCUCCCAACGGAUGUACCAGUG	\	RNA
miR-150 (quadruple base mismatch)	UCUCCCAACGGAAGUACCAGUG	\	RNA
miR-150 (five base mismatch)	UCUCCCAAGGGAAGUACCAGUG	\	RNA
miR-150 (six base mismatch)	UCUCCCAAGGGAACUACCAGUG	\	RNA

Supplementary Table 8. DNA/RNA sequences and modifications for fluorescence-based EnCODE performance validation.

Name	Sequences (5'-3')	Modification	Nucleic Acid Types
Biotin-Capture DNA 150	AATCGTGCTTGCCTGATAACCT	5’-biotin	DNA
Biotin-Capture DNA 223	AAATGCATATAGCACGGCAGAA	5’-biotin	DNA
miR-150	UCUCCCAACCCUUGUACCAGUG	\	RNA
miR-223	UGUCAGUUUGUCAAAUACCCCA	\	RNA
Tem 150	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 223	ACAAACTGACAATCAAGTAATCCAGA ATCAAGTAATCCAGAATTCTGCCGTG CTATATGCATATGGGGTATTTG	5’-PO4	DNA
Barcode 150	GTGTACTCGCCAA	3’-Cy3	DNA
Barcode 223	TCAAGTAATCCAG	3’-Fam	DNA

Supplementary Table 9. DNA/RNA sequences and modifications for multi-miRNA detection based on EnCODE.

Name	Sequences (5'-3')	Modification	Nucleic Acid Types
Biotin-Capture DNA	AATCGTGCTTGCCTGATAACCT	5’-biotin	DNA
miR-150	UCUCCCAACCCUUGUACCAGUG	\	RNA
miR-145	GUCCAGUUUUCCCAGGAAUCCCU	\	RNA
miR-636	UGUGCUUGCUCGUCCCGCCCGCA	\	RNA
miR-223	UGUCAGUUUGUCAAAUACCCCA	\	RNA
Tem 150	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 145	GGAAAACTGGACTCAAGTAATCCAGA ATCAAGTAATCCAGAAAGGTTATCAG GCAAGCACGAAAGGGATTCCTG	5’-PO4	DNA
Tem 636	CGAGCAAGCACAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAATGCGGGCGGGA	5’-PO4	DNA
Tem 223	ACAAACTGACAATCAAGTAATCCAGA ATCAAGTAATCCAGAAAGGTTATCAG GCAAGCACGAATGGGGTATTTG	5’-PO4	DNA
Barcode HRP	GTGTACTCGCCAA	3’-DBCO	DNA
Barcode AP	TCAAGTAATCCAG	3’-DBCO	DNA

Supplementary Table 10. DNA/RNA sequences and modifications for weighted calculation based on EnCODE.

Name	Sequences (5'-3')	Modification	Nucleic Acid Types
Biotin-Capture DNA	AATCGTGCTTGCCTGATAACCT	5’-biotin	DNA
miR-150	UCUCCCAACCCUUGUACCAGUG	\	RNA
miR-145	GUCCAGUUUUCCCAGGAAUCCCU	\	RNA
miR-636	UGUGCUUGCUCGUCCCGCCCGCA	\	RNA
miR-223	UGUCAGUUUGUCAAAUACCCCA	\	RNA
Tem 150-T0 (methods: template ratio weighted)	GGGTTGGGAGAAGTCAGTCTGTCAA AAGTCAGTCTGTCAAAAAGGTTATCA GGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 150-T1 (methods: template ratio weighted)	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 145-T0 (methods: template ratio weighted)	GGAAAACTGGACGTCAGTCTGTCAA AAGTCAGTCTGTCAAAAAGGTTATCA GGCAAGCACGAAAGGGATTCCTG	5’-PO4	DNA
Tem 145-T1 (methods: template ratio weighted)	GGAAAACTGGACTCAAGTAATCCAGA ATCAAGTAATCCAGAAAGGTTATCAG GCAAGCACGAAAGGGATTCCTG	5’-PO4	DNA
Tem 636-T0 (methods: template ratio weighted)	CGAGCAAGCACAGTCAGTCTGTCAA AAGTCAGTCTGTCAAAAAGGTTATCA GGCAAGCACGAATGCGGGCGGGA	5’-PO4	DNA
Tem 636-T1 (methods: template ratio weighted)	CGAGCAAGCACAGTGTACTCGCCAA AAGTGTACTCGCCAAAAAGGTTATCA GGCAAGCACGAATGCGGGCGGGA	5’-PO4	DNA
Tem 223-T0 (methods: template ratio weighted)	ACAAACTGACAAGTCAGTCTGTCAA AAGTCAGTCTGTCAAAAAGGTTATCA GGCAAGCACGAATGGGGTATTTG	5’-PO4	DNA
Tem 223-T1 (methods: template ratio weighted)	ACAAACTGACAATCAAGTAATCCAGA ATCAAGTAATCCAGAAAGGTTATCAG GCAAGCACGAATGGGGTATTTG	5’-PO4	DNA
Tem 150-T0 (methods: template type weighted)	GGGTTGGGAGAAGTCAGTCTGTCAA AAGTCAGTCTGTCAAAAGTCAGTCTG TCAAAAGTCAGTCTGTCAAAGGTTAT CAGGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 150-T1 (methods: template type weighted)	GGGTTGGGAGAAGTGTACTCGCCAA AAGTCAGTCTGTCAAAAGTCAGTCTG TCAAAAGTCAGTCTGTCAAAGGTTAT CAGGCAAGCACGAACACTGGTACAA	5’-PO4	DNA
Tem 150-T2	GGGTTGGGAGAAGTGTACTCGCCAA	5’-PO4	DNA

(methods: template type weighted)	AAGTGTACTCGCCAAAAGTCAGTCTG TCAAAAGTCAGTCTGTCAAAGGTTAT CAGGCAAGCACGAACACTGGTACAA		
Tem 150-T3 (methods: template type weighted)	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAGTGTACTCG CCAAAAGTCAGTCTGTCAAAGGTTAT CAGGCAAGCACGAACACTGGTACAA	5'-PO4	DNA
Tem 150-T4 (methods: template type weighted)	GGGTTGGGAGAAGTGTACTCGCCAA AAGTGTACTCGCCAAAAGTGTACTCG CCAAAAGTGTACTCGCCAAAGGTTAT CAGGCAAGCACGAACACTGGTACAA	5'-PO4	DNA
Barcode HRP	GTGTACTCGCCAA	3'-DBCO	DNA
Barcode AP	TCAAGTAATCCAG	3'-DBCO	DNA

Supplementary Table 11. DNA sequences and modifications for clinical sample detection based on EnCODE.

Name	Sequences (5'-3')	Modification
miR-154-5p Primer-F	TAGGTTATCCGTGTTGCCTTCG	\
miR-629-5p Primer-F	GTGGGTTTACGTTGGGAGAACT	\
miR-99a-5p Primer-F	AACCCGTAGATCCGATCTTGTG	\
miR-5006-5p Primer-F	GTTGCCAGGGCAGGAGGTGGAA	\
miR-575 Primer-F	TCTGAGCCAGTTGGACAGGAGC	\
Tem 154-5p-T1	CGGATAACCTAAGTGTACTCGCCAAAAGTGTAC TCGCCAAAAAGGTTATCAGGCAAGCACGAACG AAGGCAACA	5’-PO4
Tem 154-5p-T0	CGGATAACCTAAGTCAGTCTGTCAAAAGTCAGT CTGTCAAAAAGGTTATCAGGCAAGCACGAACG AAGGCAACA	5’-PO4
Tem 629-5p-T1	CGTAAACCCACAGTGTACTCGCCAAAAGTGTA CTCGCCAAAAAGGTTATCAGGCAAGCACGAAA GTTCTCCCAA	5’-PO4
Tem 629-5p-T0	CGTAAACCCACAGTCAGTCTGTCAAAAGTCAG TCTGTCAAAAAGGTTATCAGGCAAGCACGAAA GTTCTCCCAA	5’-PO4
Tem 99a-5p-T1	ATCTACGGGTTAGTGTACTCGCCAAAAGTGTAC TCGCCAAAAAGGTTATCAGGCAAGCACGAACA CAAGATCGG	5’-PO4
Tem 99a-5p-T0	ATCTACGGGTTAGTCAGTCTGTCAAAAGTCAGT CTGTCAAAAAGGTTATCAGGCAAGCACGAACA CAAGATCGG	5’-PO4
Tem 5006-5p-T1	GCCCTGGCAACATCAAGTAATCCAGAATCAAGT AATCCAGAAAGGTTATCAGGCAAGCACGAATT CCACCTCCT	5’-PO4
Tem 5006-5p-T0	GCCCTGGCAACAGTCAGTCTGTCAAAAGTCAG TCTGTCAAAAAGGTTATCAGGCAAGCACGAATT CCACCTCCT	5’-PO4
Tem 575-T1	ACTGGCTCAGAATCAAGTAATCCAGAATCAAGT AATCCAGAAAGGTTATCAGGCAAGCACGAAGC TCCTGTCCA	5’-PO4
Tem 575-T0	ACTGGCTCAGAAGTCAGTCTGTCAAAAGTCAG TCTGTCAAAAAGGTTATCAGGCAAGCACGAAG CTCCTGTCCA	5’-PO4
Barcode HRP	GTGTACTCGCCAA	3’-DBCO
Barcode AP	TCAAGTAATCCAG	3’-DBCO
Biotin-Capture DNA	AATCGTGCTTGCCTGATAACCT	5’-biotin

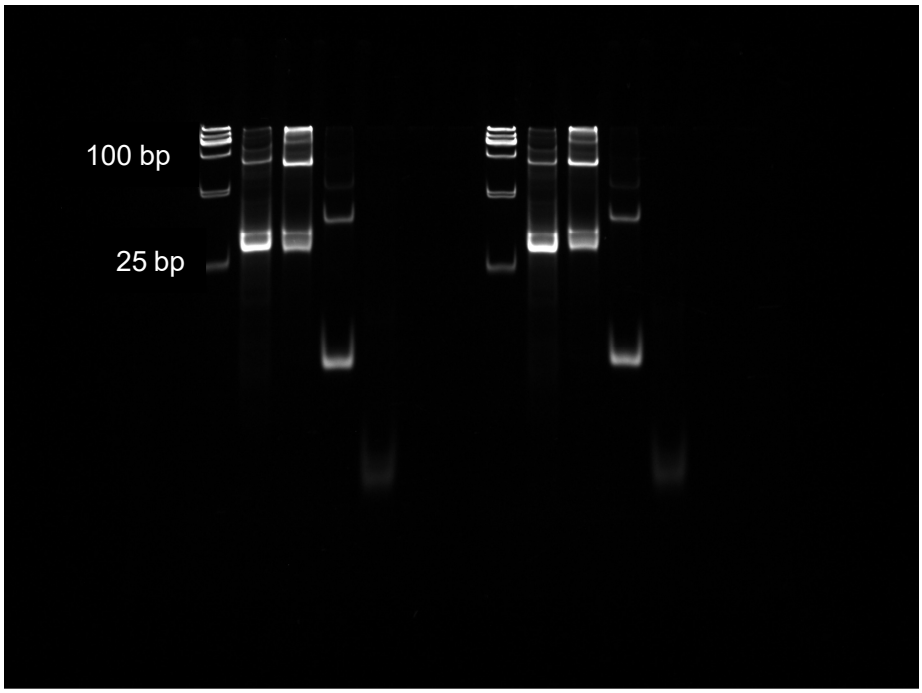
Computer Code

```
setwd("folder path ")
rm(list = ls())
options(stringsAsFactors=F)

library(limma)

logFoldChange=1
adjustP=0.05
rt=read.csv("file name")
rt=as.matrix(rt)
rownames(rt)=rt[,1]
exp=rt[,2:ncol(rt)]
dimnames=list(rownames(exp),colnames(exp))
rt=matrix(as.numeric(as.matrix(exp)),nrow=nrow(exp),dimnames=dimnames)

#group_list[group_list=='1'] <- 'PDDA'
#group_list[group_list=='0'] <- 'Control'
modType <- c(rep("PC",851),rep("NC",1972))
head(modType)
design <- model.matrix(~0+factor(modType))
colnames(design) <- levels(factor(modType))
rownames(design) <- colnames(rt)


fit <- lmFit(rt,design)
cont.matrix<-makeContrasts(paste0(unique(modType),collapse = "-"),levels=design)
fit2 <- contrasts.fit(fit, cont.matrix)
fit2 <- eBayes(fit2)
allDiff=topTable(fit2,adjust='fdr',number=200000)
diffSig <- allDiff[with(allDiff, (abs(logFC)>logFoldChange & adj.P.Val < adjustP )), ]
diffUp <- allDiff[with(allDiff, (logFC>logFoldChange & adj.P.Val < adjustP )), ]
diffDown <- allDiff[with(allDiff, (logFC<(-logFoldChange) & adj.P.Val < adjustP )), ]
write.table(diffUp,file="up.xls",sep="\t",quote=F,row.names=T)
write.table(diffDown,file="down.xls",sep="\t",quote=F,row.names=T)
write.table(allDiff,file="all.xls",sep="\t",quote=F,row.names=T)


group<-modType
group<-as.data.frame(group)
rownames(group)<-colnames(rt)
heat<-rt[rownames(rt)                                                                %in%
c(head(rownames(subset(diffSig,diffSig$logFC>0)),20),head(rownames(subset(diffSig,diffSig$logFC<
```

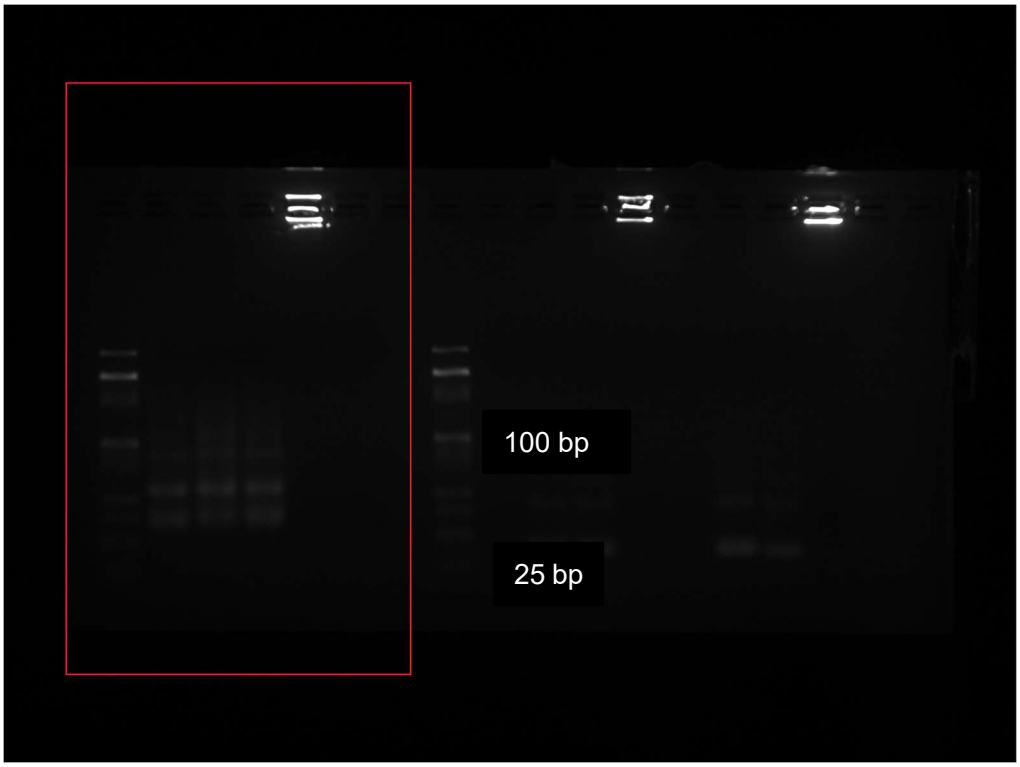
```
0)),20)),]
library(pheatmap)
x <- t(scale(t(heat)))
pheatmap(x,annotation_col=group)


png(file="1.png",width = 600,height = 600)

dev.off()
```

Uncropped scans of all gels in Supplementary Figures

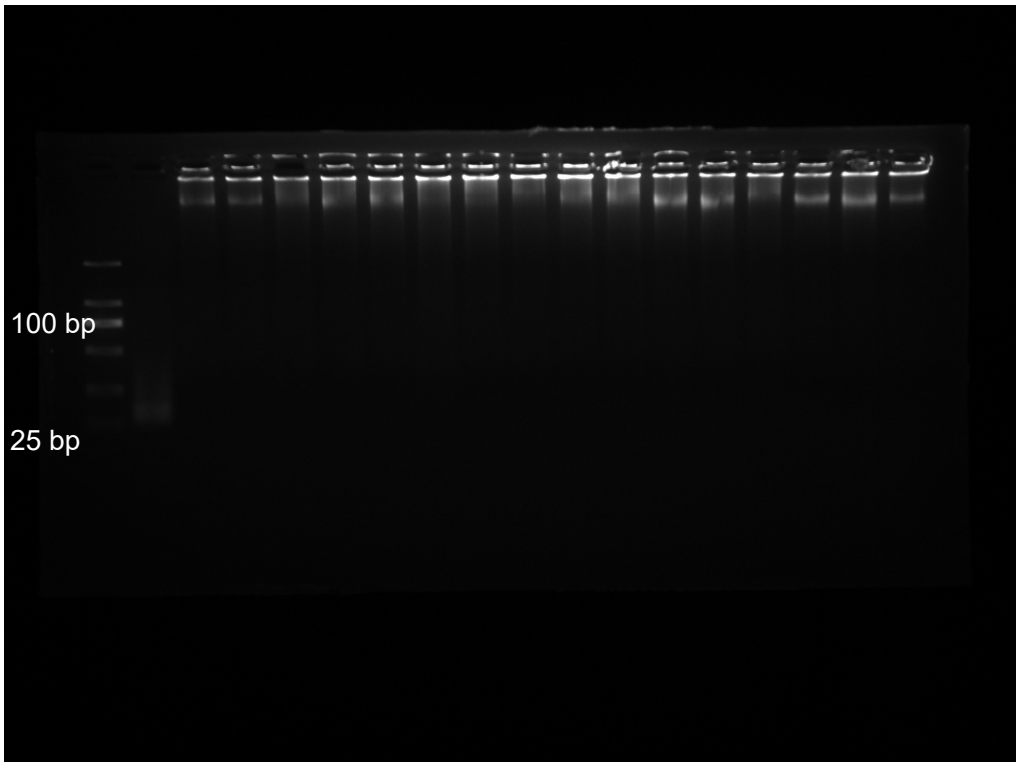
Supplementary Fig. 5a:



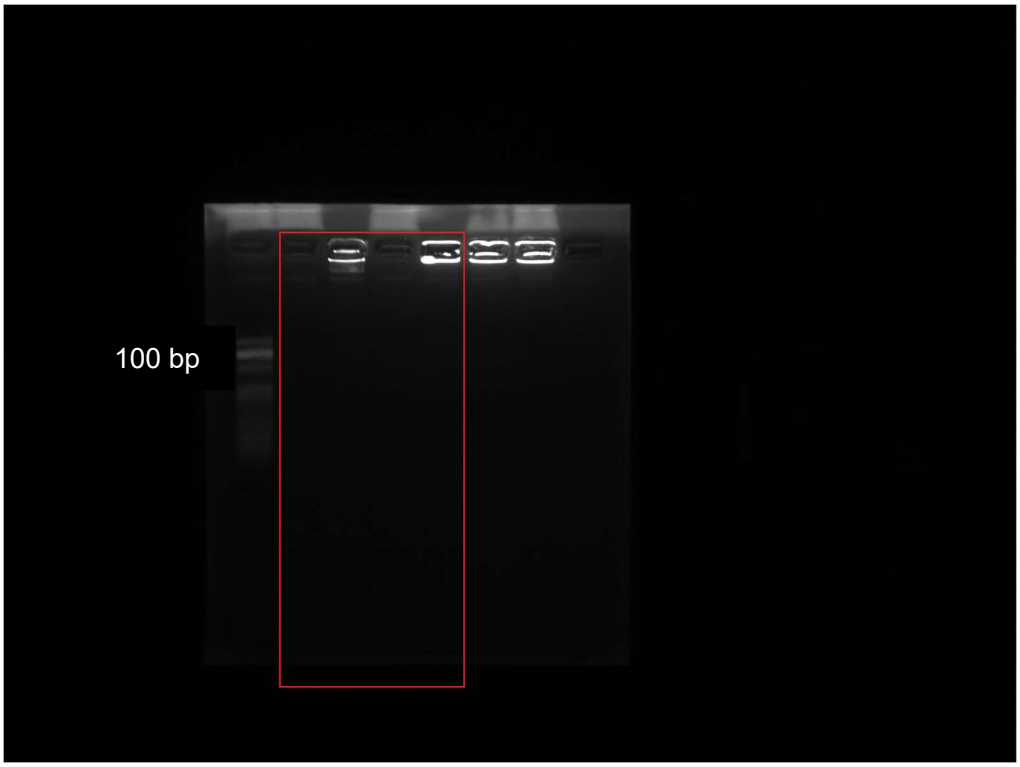
Supplementary Fig. 5b:



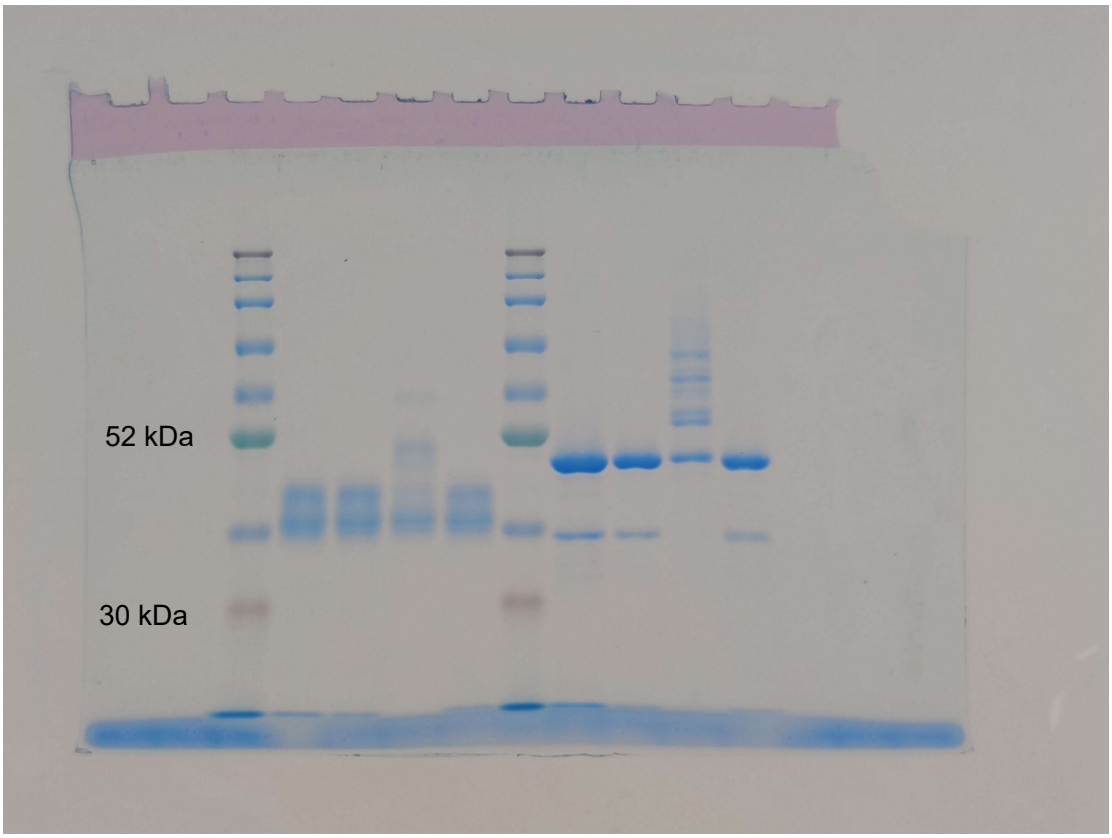
Supplementary Fig. 5d:



Supplementary Fig. 5e:



Supplementary Fig. 6b:



Supplementary Fig. 12d: