nature portfolio

Peer Review File

Enzymatic Colorimetric Encoding-Based Digital Medicine for Pancreatic Cancer Diagnosis

Corresponding Author: Professor Xiaoli Zhu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors propose and systematically validate a digital medicine platform, EnCODE, based on enzyme-driven colorimetric encoding. Using multiple miRNAs as an example, the platform maps multidimensional biomarkers into 2D/3D color signals via programmable DNA reaction and enzyme-catalyzed chromogenesis, and establishes a rigorous encode–decode framework with continuous weighting. This enables weighted computation of multiple targets and intuitive readout within a single tube. Diagnostic performance is also evaluated on clinical samples, demonstrating potential for POCT applications and multi-disease classification. Overall, the study shows strong innovation and engineering feasibility with a largely complete data chain. However, several methodological and translational aspects require further strengthening. Before acceptance, we recommend addressing the following:
1) The current encoding/decoding relies on two premises: (i) the mixed spectrum ≈ a proportional linear combination of individual enzyme spectra minus the blank; and (ii) RGB signals are linearly additive. However, validation to date covers only a limited concentration range under ideal background conditions. We recommend extending tests to both high and low extremes and quantitatively determining the conditions under which these linearity assumptions hold and where they break down—i.e., clearly defining the applicability limits of spectral linearity and RGB additivity.
2) The current weighting strategy controls weights via the Tem1/Tem0 ratio. In essence, this approach hinges on the fraction of binding sites on RCA products that are actually accessible. That accessibility can be affected by multiple factors—e.g., template folding, variability in RCA fragment length distribution, and barcode crowding that introduces steric hindrance. Consequently, non-negligible drift may arise across batches and even between wells within the same batch. In parallel, we recommend a quantitative comparison of the three weighting schemes (dynamic range, sensitivity, robustness, and CV), providing statistical evidence for the superiority of scheme (i) rather than relying on qualitative claims.
3) The manuscript states “2 mL of serum with a report within 7 hours,” but it does not specify per-step durations, which steps can run in parallel, the per-run/sample throughput, or the degree of automation.
4) It is necessary to clarify how much hemolysis and common drugs/anticoagulants affect the HRP/AP reactions and the optical readouts. Based on these results, develop actionable rules that specify when to reject a sample and when—and how—to apply correction.
5) The authors use NHS-PEG4-N3/DBCO click chemistry to conjugate enzymes to DNA and claim that “enzyme activity is unaffected.” However, key quantitative evidence is still lacking to substantiate this claim—particularly the distribution of modification degrees, the enzyme load per RCA product, the conjugation yield, and the specific activity per barcode. These parameters directly determine the metrological stability of the miRNA-to-enzyme signal transfer.

Reviewer #2

(Remarks to the Author)
This study proposes an innovative enzymatic colorimetric encoding (EnCODE) platform that enables weighted calculation of multiple miRNA biomarkers within a single tube and translates the results into a visual color signal. The concept is novel, and the technology demonstrates diagnostic performance in pancreatic cancer comparable to qRT-PCR (AUC = 0.97), suggesting both academic value and potential for clinical translation. However, the manuscript requires major revisions in several critical areas

Major Revisions

1. The transformation of color signals into risk scores may be influenced by multiple confounding factors (e.g., sample handling time, batch effects, environmental temperature and humidity). A detailed description of the calibration strategy, including whether standardized reference materials were used, as well as data on robustness, is required.
2. The rationale behind selecting the five miRNA markers is insufficient. The manuscript would benefit from a detailed description of the selection pipeline, supporting evidence from the literature, and a discussion of their biological roles in pancreatic cancer development.
3. The process for determining the coefficients in the weighted calculation formula is unclear. The authors should describe the optimization strategy in detail (e.g., machine learning algorithms, statistical modeling), and provide justification for the chosen approach.
4. A direct comparison with existing point-of-care testing (POCT) platforms is missing. To highlight the novelty and translational value of EnCODE, please provide comparative data with at least two commercially available rapid detection methods, including cost-effectiveness and operational complexity assessments.
5. The current study includes only 32 samples (12 pancreatic cancer patients), which is insufficient to draw reliable conclusions. The authors are strongly encouraged to increase the sample size to a statistically meaningful scale and provide a clear rationale with sample size calculations.
6. An independent external validation cohort is missing. To establish the robustness and generalizability of the EnCODE platform, data from additional medical centers and different populations should be included.
7. The control group design is suboptimal, as it does not incorporate benign pancreatic diseases (e.g., chronic pancreatitis). Including such cases is essential for accurately evaluating diagnostic specificity.
8. Critical clinical information (e.g., disease stage, pathological type) is lacking for the enrolled patients. Please provide complete baseline clinical characteristics.
9. The GSE163031 dataset used for validation contains too few samples, raising concerns about potential overfitting. A larger external dataset should be used for validation, or appropriate overfitting tests should be presented.

Minor Revisions
10. The manuscript should include a detailed description of how the concentrations of EnCODE components (primers, templates, enzymes) were optimized and determined.
11. Please clarify the optimization strategy for the coefficients in the weighted calculation formula, and indicate whether alternative mathematical models were tested and how their performance compared.
12. Additional validation data demonstrating the linear relationship between color signals and miRNA concentrations, including the linear detection range, should be provided.
13. The effects of sample storage conditions (e.g., temperature, storage time) on assay performance should be systematically assessed.
14. Please provide a quantitative analysis of how common interfering factors (e.g., hemolysis, lipemia) affect assay accuracy.

15. Variability in subjective interpretation of colorimetric results should be evaluated, and strategies to improve consistency should be proposed.
16. Finally, please describe the pathway and technical feasibility of smartphone integration with the EnCODE platform.

Reviewer #3

(Remarks to the Author)
Peer Review Report
This manuscript presents EnCODE, an enzymatic colorimetric encoding system for detecting multiple miRNA biomarkers in clinical samples. While the core technical concept demonstrates some merit, substantial concerns limit the work's current contribution and reliability.
The authors report several noteworthy results, including the development of a dual-enzyme system that mathematically encodes biomarker concentrations into interpretable color patterns, demonstration of linear relationships between enzyme concentrations and spectral outputs, and extension to three-dimensional encoding. In a clinical study of 32 samples, they achieved reported sensitivity of 100% and overall accuracy of 91% for pancreatic cancer detection. However, these results require critical examination.
Regarding significance to the field, the work makes a meaningful contribution to colorimetric biosensing and multiplexed nucleic acid detection. The mathematical encoding framework represents a novel approach - while colorimetric detection exists widely, the systematic use of enzymatic color mixing as a mathematically rigorous information encoding system with reversible decoding capabilities is innovative. This differs substantially from conventional colorimetric approaches that typically provide simple positive/negative or semi-quantitative readouts. However, the clinical diagnostic performance does not clearly exceed established miRNA detection methods, and without comparative studies, the practical advantages over qRT-PCR or digital PCR remain unclear. The clinical application, while demonstrating feasibility, requires more robust validation to assess its true diagnostic value.
The work does not adequately support its major conclusions and claims. The 32-sample clinical validation is insufficient for reliable diagnostic performance estimates, requiring independent validation cohorts and multi-site studies. Claims of "Polaroid-like simplicity" contradict the multi-step, 7-hour protocol described, and no evidence supports visual interpretation by non-specialists. Assertions of superior performance over established methods lack supporting comparative data, and the biomarker selection process from existing datasets may involve circularity between discovery and validation.
Several significant flaws in data analysis, interpretation, and conclusions prohibit publication in the current form. Statistical issues include inadequate sample size for diagnostic performance claims, wide confidence intervals indicating high uncertainty, potential overfitting with five variables and only twelve positive cases, and unclear validation methodology.

Analytical concerns encompass missing validation of linearity under real-world conditions, insufficient assessment of matrix effects, limited stability data, and no inter-operator validation. The interpretation problems include claims exceeding supporting evidence, premature clinical readiness assertions, and oversimplified presentation of the complex protocol.
The methodology falls below expected standards for diagnostic test evaluation. The clinical study lacks proper power analysis and adequate controls, biomarker selection may involve methodological circularity, and critical validation studies for stability, reproducibility, and specificity are missing. Environmental variables affecting colorimetric detection are insufficiently controlled, and potential confounding factors receive limited assessment. While the basic chemistry appears sound, the clinical validation methodology is inadequate for the diagnostic claims presented.
Critical details necessary for reproduction are missing throughout the methods section. These include enzyme standardization and quality control procedures, environmental control requirements, specific protocols for visual interpretation, statistical analysis software parameters, detailed preparation protocols for DNA templates and conjugates, and storage requirements for reagents. Importantly, despite stating that code is available "upon reasonable request," no data analysis scripts, algorithms for spectral decoding, or computational workflows are provided or deposited in public repositories. This lack of computational transparency significantly hampers reproducibility, particularly given the mathematical encoding/decoding framework central to the method. Although supplementary methods provide some additional detail, key aspects remain insufficiently described for independent reproduction.
The manuscript requires major revision before publication consideration. The core technical approach has merit as a proof-of-concept study, but clinical claims are premature and inadequately supported. The authors should reframe this as a methods development study rather than a clinical diagnostic platform, conduct properly powered clinical validation with independent cohorts, provide comprehensive reproducibility and stability data, include direct comparisons with established detection methods, make all analysis code and algorithms publicly available, and correct the numerous presentation errors throughout the manuscript. Additionally, the pervasive spelling and grammatical errors, including technical terms in figures, require professional editing to meet publication standards.
With appropriate scope and validation, this work could potentially contribute to the biosensing literature. However, substantial additional evidence and methodological improvements are needed to support the current conclusions and clinical claims.

Minor Issues
The manuscript also needs professional editing - there are numerous spelling errors and grammatical mistakes throughout, including in figure labels (e.g., "spectural" instead of "spectral", "competitive binging" instead of "binding").

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I think the authors have fully addressed my previous concerns. I have no further comments.

Reviewer #2

(Remarks to the Author)
The authors did sufficient job to address my critics.

Reviewer #3

(Remarks to the Author)
The authors have substantially improved their manuscript which now meets the standard. This is a very good and novel study which should be interested for a broad readership.

Response to Reviewers

Dear Editor and Reviewers,

We are sincerely grateful for the thoughtful feedback on our manuscript titled "Enzymatic Colorimetric Encoding-Based Digital Medicine for Pancreatic Cancer Diagnosis" (Manuscript ID: NCOMMS-25-68381). The reviewers' comments have not only helped us in enhancing the clarity and rigor of this work but also provided valuable perspectives that will inspire our future research directions. We have carefully addressed all comments, as outlined in the point-by-point responses below, with corresponding revisions highlighted in red font within the revised manuscript.

Reviewer #1:

The authors propose and systematically validate a digital medicine platform, EnCODE, based on enzyme-driven colorimetric encoding. Using multiple miRNAs as an example, the platform maps multidimensional biomarkers into 2D/3D color signals via programmable DNA reaction and enzyme-catalyzed chromogenesis, and establishes a rigorous encode–decode framework with continuous weighting. This enables weighted computation of multiple targets and intuitive readout within a single tube. Diagnostic performance is also evaluated on clinical samples, demonstrating potential for POCT applications and multi-disease classification. Overall, the study shows strong innovation and engineering feasibility with a largely complete data chain. However, several methodological and translational aspects require further strengthening. Before acceptance, we recommend addressing the following:

Response:

We sincerely thank you for your thoughtful and encouraging comments on our work.

1) The current encoding/decoding relies on two premises: (i) the mixed spectrum $\approx$ a proportional linear combination of individual enzyme spectra minus the blank; and (ii) RGB signals are linearly additive. However, validation to date covers only a limited concentration range under ideal background conditions. We recommend extending tests to both high and low extremes and quantitatively determining the conditions under which these linearity assumptions hold and where they break down—i.e., clearly defining the applicability limits of spectral linearity and RGB additivity.

Response:

Thank you for this important suggestion. We have now extended the linearity tests to a broader concentration range and explicitly defined the operating window where spectral linearity and RGB additivity hold, as well as the conditions where deviations arise.

Under the standard 30 min development time, HRP–ABTS achieved an LOD of 34.19 pM with excellent linearity from 68.38 pM to 2 nM ($R^2 > 0.99$), and AP–pNPP achieved an LOD of 14.88 pM with linearity from 31.24 pM to 2 nM ($R^2 > 0.99$). At enzyme concentrations > 2 nM, absorbance values exceeded the measurement range of the plate reader (TECAN Infinite 200 PRO), which therefore defines a practical upper bound for linear spectral quantification. For the RGB-based

readout, HRP–ABTS exhibited an LOD of 68 pM with a linear range of 135 pM–2 nM ($R^2 > 0.98$), whereas AP–pNPP exhibited an LOD of 236 pM with a linear range of 430 pM–2 nM ($R^2 > 0.98$). Notably, at enzyme concentrations > 2 nM, RGB signals showed pronounced dynamic-range compression (near saturation), limiting quantitative performance (**Supplementary Figure 2**)

Extending the chromogenic reaction to 1 h improved both spectral and RGB sensitivity, providing a tunable strategy to match the expected concentration range of different sample types. For RGB readout, the LOD/LOQ reached 8.87/22.58 pM for HRP–ABTS and 8.81/19.41 pM for AP–pNPP. For spectral readout, the LOD/LOQ reached 5.54/12.25 pM for HRP–ABTS and 8.87/22.58 pM for AP–pNPP (**Supplementary Figure 3**). Accordingly, for low-abundance targets, an appropriately extended development time can be used to enhance sensitivity.

Finally, we directly assessed spectral and RGB additivity in enzyme mixtures. Spectral additivity holds when each component concentration lies within the linear absorbance regime and the mixed spectrum remains below the instrument saturation limit. Similarly, RGB additivity holds when each component lies within the linear RGB regime and the recorded RGB values remain within the non-compressed dynamic range (empirically 30–240 in 8-bit RGB acquisition), outside of which non-linear compression becomes apparent (**Supplementary Figure 4**). (**Line 143-154, Line 168-172, Supplementary Figure 2, Supplementary Figure 3 and Supplementary Figure 4)**

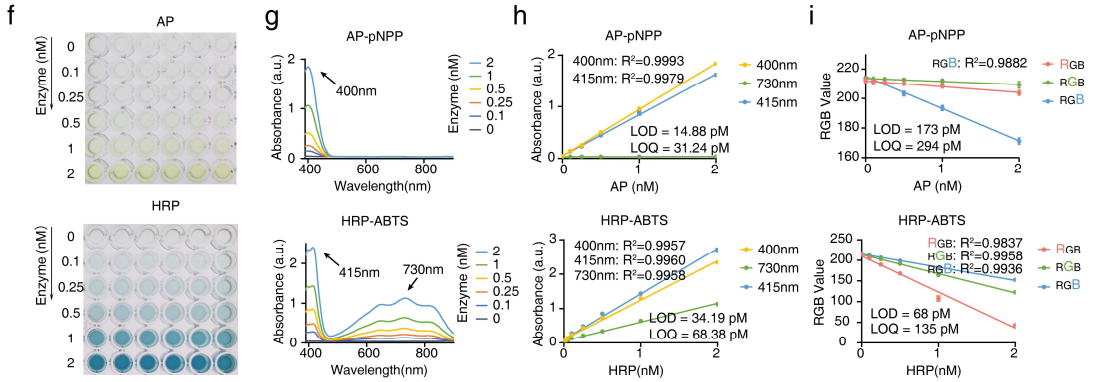
Revised Main Text (Line 143-154):

Additionally, we validated the enzymatic colorimetric performance of AP and HRP, confirming a well-defined mathematical mapping between enzyme concentration and colorimetric output. Both enzymes showed stable catalytic behavior under the selected reaction conditions (**Supplementary Figure 2f**). Across the validated dynamic range, enzyme concentration exhibited positive linear relationships with both absorbance and RGB features, supporting robust enzyme-driven color modulation (**Supplementary Figure 2g–i**). For the spectral readout, linear quantification was maintained from 68 pM to 2 nM, whereas for the RGB-based readout, the linear range was approximately 294 pM to 2 nM. For lower-abundance samples, assay sensitivity can be improved by moderately extending the chromogenic development time. Extending the color-development reaction to 1 h improved low-end sensitivity, reducing the spectral LOD to 5.54 pM (HRP–ABTS) and 1.6 pM (AP–pNPP), and the RGB-based LOD to 8.87 pM (HRP–ABTS) and 8.81 pM (AP–pNPP) (Supplementary Figure 3). (**Supplementary Figure 3**).

Revised Main Text (Line 168-172):

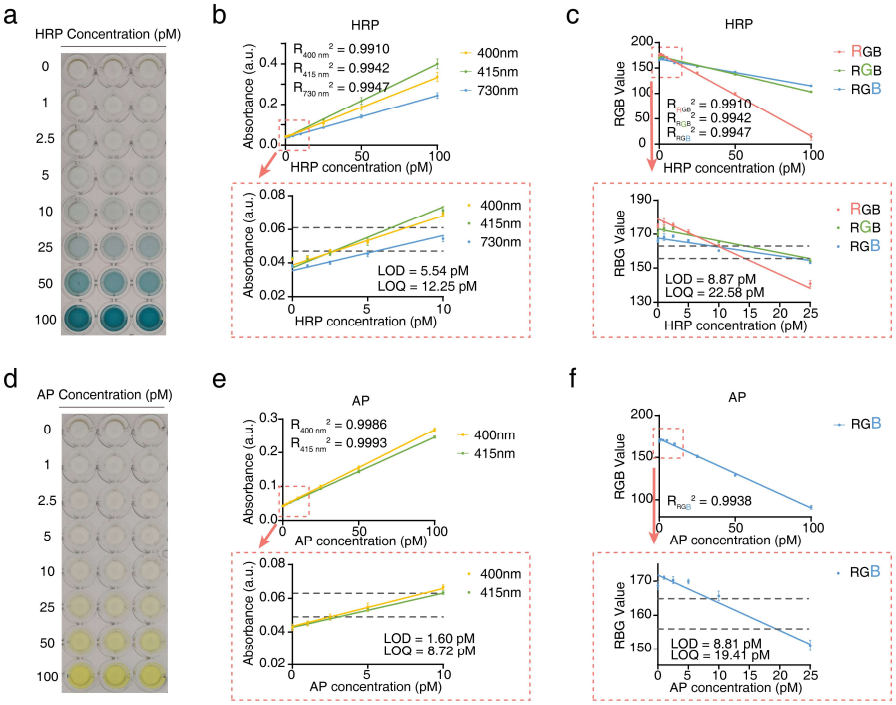
These quantitative additivity relationships, validated within the experimentally defined linear regime, established the foundation for accurate encoding and decoding of enzymatic colorimetric signals. This additivity was further confirmed for mixed-enzyme reactions at ultralow concentrations (**Supplementary Figure 4**).

Revised Supplementary Figure 2:



Supplementary Figure 2. (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).

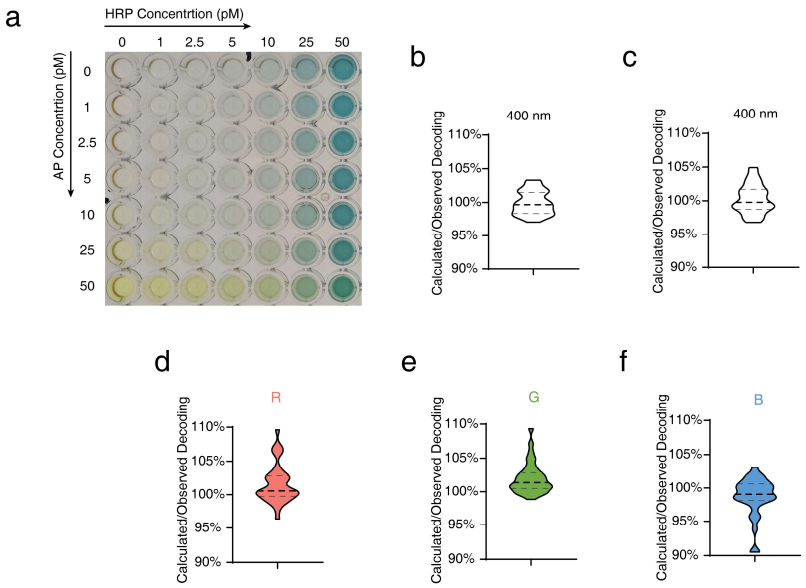
Revised Supplementary Figure 3:



Supplementary Figure 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).

Revised Supplementary Figure 4:



Supplementary Figure 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). (c) Ratio of predicted-to-measured absorbance at 415 nm for each well in panel (a). (d) Ratio of predicted-to-measured R-channel values for each well in panel (a). (e) Ratio of predicted-to-measured G-channel values for each well in panel (a). (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.

2) The current weighting strategy controls weights via the Tem1/Tem0 ratio. In essence, this approach hinges on the fraction of binding sites on RCA products that are actually accessible. That accessibility can be affected by multiple factors — e.g., template folding, variability in RCA fragment length distribution, and barcode crowding that introduces steric hindrance. Consequently, non-negligible drift may arise across batches and even between wells within the same batch. In parallel, we recommend a quantitative comparison of the three weighting schemes (dynamic range, sensitivity, robustness, and CV), providing statistical evidence for the superiority of scheme (i) rather than relying on qualitative claims.

Response:

Thank you for this important comment on weighting robustness. We performed controlled experiments to quantify both within-batch (well-to-well) and between-batch variability under fixed target concentrations and identical processing conditions. Statistical comparisons across independent batches confirmed superior sensitivity and weighting responsiveness for scheme (i).

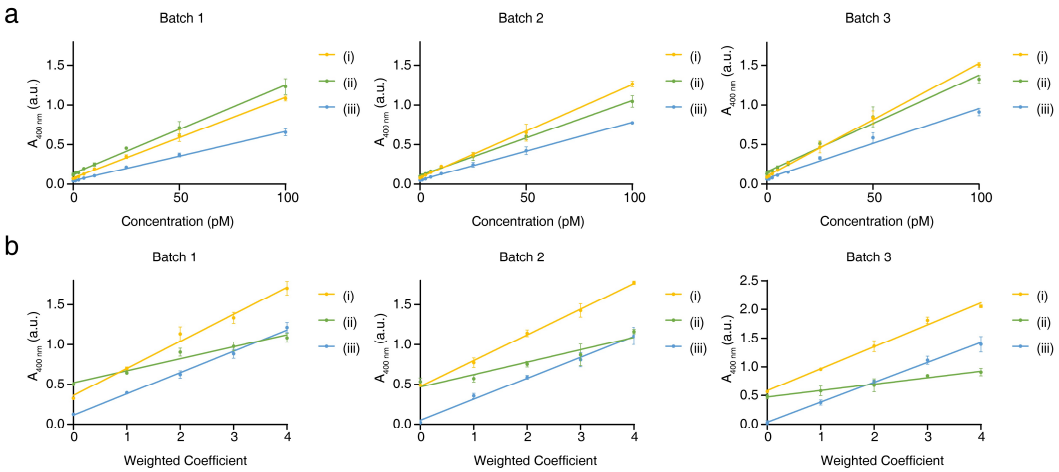
In parallel, we conducted a quantitative comparison of the three weighting schemes using analytical metrics including dynamic range, sensitivity (slope), linearity (R^2), and robustness (CV%).

At a representative weight coefficient (0.5), all three schemes exhibited comparable limits of detection (LOD), indicating that weighting does not compromise the fundamental detection capability. However, scheme (iii) consistently showed lower sensitivity (smaller slope), resulting in reduced quantitative resolution. We further assessed the responsiveness of each scheme to changes in the weight coefficient at a fixed target concentration. Scheme (i) displayed the largest and most monotonic output change across the weighting range, whereas scheme (ii) exhibited the weakest response, indicating reduced controllability. Importantly, scheme (ii) supports only discrete weighting. Well-to-well variability remained low for all schemes (CV% < 5%), and batch-to-batch variability was within an acceptable range (CV% < 15%). These results provide quantitative support for selecting scheme (i) as the default strategy for weighting control, and the corresponding data have been added as Supplementary Figure 11 and Supplementary Table 2-3 in the revised manuscript. **(Line 352-357, Supplementary Figure 11, Supplementary Table 2 and Supplementary Table 3)**

Revised Main Text (Line 352-357):

However, scheme (ii) supports only discrete weighting, whereas scheme (iii) showed a weaker target concentration–response, potentially due to the unlabeled DNA barcodes—experiencing less steric hindrance—can more effectively compete with RCA products for binding. Ultimately, scheme (i) was selected for subsequent weighted calculations, as it exhibited the best multi-batch performance with minimal variability (**Supplementary Figure 11 and Supplementary Tables 2–3**).

Revised Supplementary Figure 11:



Supplementary Figure 11. Multi-batch performance evaluation of three weighting schemes. (a) Target concentration–response curves (absorbance) measured at a fixed weighting coefficient for the three weighting schemes. (b) Output–weight relationship measured at a fixed target concentration. Data are presented as mean ± s.d. from three independent batches (n = 3).

Revised Supplementary Table 2:

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

Metric	scheme (i):	scheme (ii): Barcode	scheme (iii):

	Tem1/Tem0 ratio control	copy-number modulation	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.

Revised Supplementary Table 3:

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.

3) The manuscript states “2 mL of serum with a report within 7 hours,” but it does not specify per-step durations, which steps can run in parallel, the per-run/sample throughput, or the degree of automation.

Response:

We thank you for this important request for clarification. To ensure full reproducibility and transparency, we have now added a detailed breakdown of assay timing, parallelizable steps, throughput capacity, and the extent of automation. The revised manuscript includes this information in the Methods section. (**Line 813-846**)

Revised Main Text (Line 813-846):

Overview of the EnCODE Assay Workflow

The entire EnCODE assay, from sample preparation to final colorimetric readout, is completed within 7 hours, with the following breakdown:

Serum/Plasma miRNA Extraction: Serum/plasma samples are processed using the miRcute Serum/Plasma miRNA Extraction Kit (TianGen Biotech, Beijing, China). This step takes approximately 1 hour. Using a refrigerated high-speed centrifuge (Eppendorf 5424), 24 serum/plasma samples can be processed per batch. Extracted miRNAs are stored in nuclease-free water at -80°C until use.

miRNA Pre-amplification: miRNAs are reverse transcribed into cDNA using the Mir-X miRNA

First-Strand Synthesis Kit (Takara Bio, Japan), followed by LATE-PCR amplification using the 2X TaqMan Fast qPCR Master Mix (Sangon Biotech, Shanghai, China). This step requires 30–40 minutes for reverse transcription and 1–1.5 hours for PCR amplification. The sample processing capacity per batch is limited to the capacity of the PCR machine, typically 96 samples per batch.

Circularization of DNA Templates: Using the pre-amplified miRNA product as a primer, DNA templates are circularized via SplintR ligase. This step takes approximately 30–40 minutes, with no theoretical limit to the number of samples per batch.

Capture Probe Coating: The capture probe is immobilized onto a streptavidin-coated enzyme plate in 2 hours, which can be done in parallel with steps 1, 2, and 3.

Template Capture: Circularized DNA templates are captured by incubating the miRNA-pre-amplified DNA-template complex with the streptavidin-coated enzyme plate. This step takes 30 minutes. The batch processing capacity is theoretically unlimited.

Rolling Circle Amplification and Signal Encoding: Phi29 polymerase is used for amplification of the circular template, while simultaneously introducing DNA-enzyme complexes for signal encoding. This step takes 30 minutes and can handle an unlimited number of samples per batch.

Enzyme-Catalyzed Colorimetric Reaction: Substrates for the enzyme-catalyzed reaction are added, producing detectable colorimetric signals. This step requires 30 minutes. While theoretically unlimited in terms of sample processing per batch, for reducing batch-to-batch variation, it is recommended to process no more than 96 samples per batch.

Signal Detection: Signal detection takes 3–5 minutes, with the capacity to process 96 samples per batch.

The total hands-on processing time is approximately 5–6 hours. With additional error allowances for calculation and operational setup, the full process from clinical sample to result output can be completed in approximately 6–7 hours. All steps in the assay can be automated with robotic arms, reducing processing time and minimizing batch-to-batch variation.

4) It is necessary to clarify how much hemolysis and common drugs/anticoagulants affect the HRP/AP reactions and the optical readouts. Based on these results, develop actionable rules that specify when to reject a sample and when—and how—to apply correction.

Response:

We sincerely thank you for raising the important issue. We have conducted targeted interference experiments to evaluate these effects and developed actionable rules for sample rejection and correction based on our findings.

Although the EnCODE workflow primarily focuses on the extraction of free miRNA from plasma/serum samples, which does not directly affect HRP/AP reactions or optical readouts, sample conditions may impact miRNA extraction efficiency and, in turn, affect the final results.

To assess potential interference, we established three controlled experimental groups using blood samples from the same source, thereby minimizing variability arising from endogenous miRNA differences across individuals. First, to evaluate anticoagulants and coagulants, blood was collected into commonly used collection tubes (e.g., heparin- or EDTA-containing tubes, and serum tubes with clot activator), and EnCODE readouts were compared across tube conditions. Second, to assess interference from pancreatic cancer therapeutics, matched aliquots of the same blood sample were supplemented with representative drugs (5-Fluorouridine, Irinotecan, Oxaliplatin, and

Gemcitabine) at the tested concentrations prior to processing. Third, to model common matrix interferences, equal volumes of Intralipid, albumin solution, or red-blood-cell lysate were added to matched aliquots to simulate lipemia, hyperproteinemia, and hemolysis, respectively.

All groups were processed through the complete EnCODE workflow, and outputs were compared against the corresponding untreated aliquot controls (**Supplementary Figure 14a**). Anticoagulants/coagulants and the tested therapeutic agents did not produce appreciable changes in EnCODE readouts, and neither lipemic nor high-protein conditions showed significant interference. In contrast, hemolysis led to a modest upward shift in measured miRNA levels, consistent with the release of endogenous miRNAs from lysed erythrocytes.

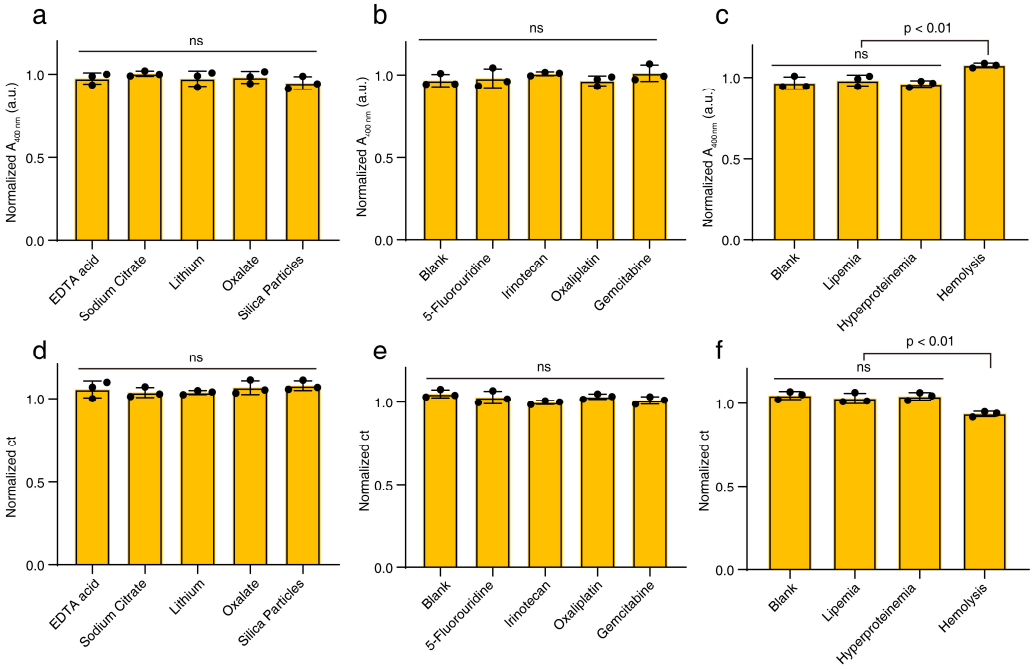
In addition, EnCODE readouts showed the same directional trends as those obtained by qRT-PCR on the matched samples, supporting that the tested sample conditions do not materially bias EnCODE-based quantification of miRNA abundance (**Supplementary Figure 14b**).

Based on these findings, we conclude that hemolytic samples can affect miRNA detection, and thus, careful evaluation of sample hemolysis should be conducted before processing. (**Line 414-430 and Supplementary Figure 14)**

Revised Main Text (Line 414-430):

To evaluate robustness under clinically relevant conditions, we assessed common pre-analytical interferences and therapeutic agents using aliquots from the same-source blood sample. We tested anticoagulants/coagulants, representative pancreatic cancer drugs (5-Fluorouridine, Irinotecan, Oxaliplatin, and Gemcitabine), and simulated matrix interferences (lipemia, high protein, and hemolysis) by adding equal volumes of Intralipid, albumin, or red-blood-cell lysate, respectively. Anticoagulants/coagulants, drugs, and lipemic/high-protein conditions showed no appreciable impact on EnCODE readouts, whereas hemolysis caused a modest upward shift consistent with erythrocyte miRNA release (**Supplementary Figure 14a-c**). Notably, EnCODE readouts followed the same directional trends as qRT-PCR on the matched samples, supporting that these sample conditions do not materially bias miRNA quantification (**Supplementary Figure 14d-f**).

Revised Supplementary Figure 14:



Supplementary Figure 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 Student's t-test, and p values are indicated.

5) The authors use NHS-PEG4-N3/DBCO click chemistry to conjugate enzymes to DNA and claim that "enzyme activity is unaffected." However, key quantitative evidence is still lacking to substantiate this claim—particularly the distribution of modification degrees, the enzyme load per RCA product, the conjugation yield, and the specific activity per barcode. These parameters directly determine the metrological stability of the miRNA-to-enzyme signal transfer.

Response:

We sincerely thank you for raising the important point. To address this concern, we have conducted new measurements and analyses.

Using SDS-PAGE, we analyzed the degree of DNA modification on the HRP and AP enzymes after click conjugation. Both enzymes exhibited a narrow distribution of 1-3 DNA molecules per enzyme, with species carrying ⩾2 DNA groups comprising less than 30% of the population. The conjugation yield was also evaluated, which was found to be approximately 70%, meeting the requirements for the EnCODE platform (**Supplementary Figure 3a-b**).

We then assessed whether click conjugation affected the enzymatic activity of HRP and AP.

We compared the enzyme activity of conjugated and unmodified enzymes using standardized catalytic assays. The conjugated enzymes showed 78% of the activity of native HRP and 84% of the activity of native AP. These results confirm that, within the operating range of the assay, the conjugation process does not significantly impact the catalytic turnover efficiency (**Supplementary Figure 3c**).

In addition, to estimate the enzyme loading per RCA product, we performed a plate-based comparison. Specifically, we quantified the colorimetric output obtained by directly capturing the enzyme–DNA conjugates on the capture strand (i.e., without RCA), and compared it with the output obtained after a 30-min RCA step, where the RCA products subsequently captured the same enzyme–DNA conjugates prior to chromogenic development. For the HRP–DNA conjugate, the RCA-enabled workflow produced an approximately 28-fold higher signal than direct capture, indicating an average HRP loading of ~28 per RCA product under the tested conditions. Using the same approach, the AP–DNA conjugate showed an approximately 26-fold signal increase after RCA, corresponding to an average AP loading of ~26 per RCA product (**Supplementary Figure 7)**.

Finally, we benchmarked EnCODE against qRT-PCR using matched samples to directly assess the metrological stability of the miRNA-to-enzyme signal transfer. Consistent with the processing of clinical samples, the target miRNAs in this comparison were subjected to the same pre-amplification step prior to EnCODE. EnCODE measurements showed strong agreement with qRT-PCR quantification across the tested range ($R^2 > 0.98$). Moreover, EnCODE readouts were highly correlated with the logarithm of the target miRNA abundance ($R^2 > 0.98$), supporting a stable and quantitatively reliable mapping from miRNA input to enzymatic colorimetric output (**Supplementary Figure 13**). **(Line 235-248, Line 406-413, Supplementary Figure 6, Supplementary Figure 7 and Supplementary Figure 13)**.

Revised Main Text (Line 235-248):

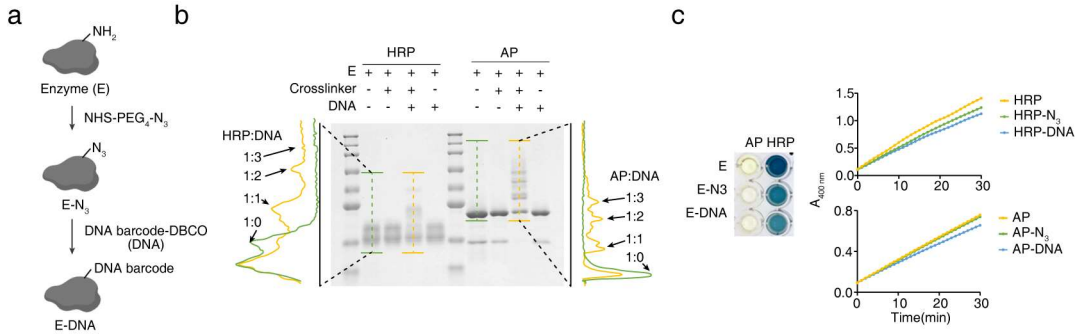
Also, cross-linking enzymes to DNA barcodes via NHS-PEG4-N3/DBCO click chemistry achieved efficient conjugation. The conjugation yield was approximately 70%, and each enzyme was modified with 1-3 DNA molecules. Enzyme activity recmained stable, with 78% of the activity of native HRP and 84% of native AP, confirming that the conjugation process did not significantly affect enzymatic performance. These results validate the chemical integrity and metrological stability of the enzyme-DNA conjugates used in EnCODE (**Supplementary Figure 6**). To further estimate the enzyme loading per RCA product, we compared colorimetric outputs obtained by directly capturing enzyme–DNA conjugates on the plate-immobilized capture strand (without RCA) versus capturing the same conjugates via RCA products after a 30-min amplification. The RCA-enabled workflow increased signal intensity by ~28-fold for HRP–DNA and ~26-fold for AP–DNA, consistent with average loadings of ~28 HRP and ~26 AP molecules per RCA product under the working conditions. These results support efficient signal amplification while maintaining a quantitatively stable miRNA-to-enzyme signal transfer (**Supplementary Figure 7**).

Revised Main Text (Line 406-413):

To further assess metrological stability in a workflow-relevant setting, we benchmarked EnCODE using miRNA samples processed through the same pre-amplification step as in the clinical workflow. Across the tested concentration series, EnCODE readouts closely followed the input miRNA levels and showed strong agreement with qRT-PCR quantification, supporting reliable

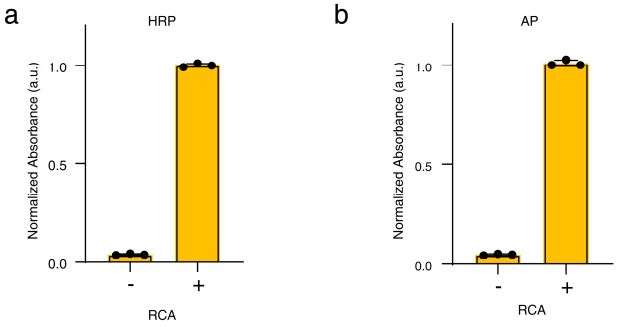
quantitative transfer from miRNA abundance to enzymatic output (**Supplementary Figure 13a–c**). In parallel, both the HRP–ABTS and AP–pNPP systems exhibited robust linear relationships between miRNA concentration and colorimetric features at the levels of absorbance and RGB values, further supporting a stable miRNA-to-enzyme signal mapping (**Supplementary Figure 13d-e**).

Revised Supplementary Figure 6:



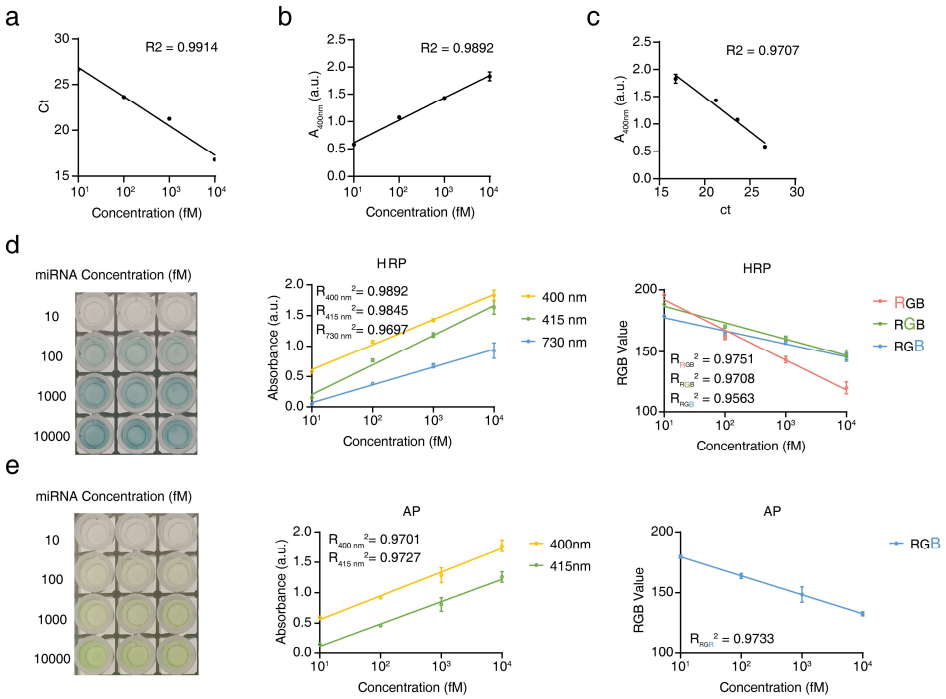
Supplementary Figure 6. Verification of the cross-linking of enzymes with DNA barcodes. (a) Schematic diagram of the cross-linking process between enzymes and DNA barcodes. (b) SDS-PAGE electrophoresis image of the successful cross-linking between AP and HRP with DNA barcodes. (c) Enzymatic kinetics analysis of the enzyme before and after its cross-linking with DNA barcodes.

Revised Supplementary Figure 7:



Supplementary Figure 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). Data are presented as the mean ± s.d. from three independent experiments (n = 3).

Revised Supplementary Figure 13:



Supplementary Figure 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) 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).

Reviewer #2:

This study proposes an innovative enzymatic colorimetric encoding (EnCODE) platform that enables weighted calculation of multiple miRNA biomarkers within a single tube and translates the results into a visual color signal. The concept is novel, and the technology demonstrates diagnostic performance in pancreatic cancer comparable to qRT-PCR (AUC = 0.97), suggesting both academic value and potential for clinical translation. However, the manuscript requires major revisions in several critical areas

Response:

We sincerely thank you for the thoughtful, balanced, and constructive evaluation of our work. We appreciate the recognition of the conceptual and methodological contribution of EnCODE, and

we have carefully revised the manuscript to address all concerns raised.

Major Revisions

1. The transformation of color signals into risk scores may be influenced by multiple confounding factors (e.g., sample handling time, batch effects, environmental temperature and humidity). A detailed description of the calibration strategy, including whether standardized reference materials were used, as well as data on robustness, is required.

Response:

We thank you for highlighting that the transformation from colorimetric signals to risk scores may be affected by pre-analytical and environmental confounders. Accordingly, we have expanded the Methods and Supplementary Information to provide a detailed description of the calibration workflow, reference materials, and robustness tests.

All sample handling and EnCODE procedures were performed under environmental chamber (25 ± 1 °C and 50 ± 5% relative humidity). For image-based readout, we additionally standardized the imaging geometry using a fixed-distance holder and employed uniform diffuse illumination (LED ring light or a backlight lightbox) to minimize shadows and spatial non-uniformity. Each assay run included internal reference standards for plate-to-plate normalization, thereby correcting potential variations in spectral response and enzyme activity.

To evaluate the impact of sample handling delays, we subjected serum samples to defined room-temperature holding times prior to miRNA extraction; the decoded outputs showed no significant drift within the tested time window (**Supplementary Figure 10c-d**). Batch robustness was assessed using multiple independent runs performed under identical environmental conditions, and the post-calibration variability remained within the expected range for enzymatic colorimetric assays (**Supplementary Figure 10a-b**). These details and associated robustness results have been added to the revised Supplementary Information. (**Line 306-310, Line 653-657, Line 741-745 and Supplementary Figure 10**)

Revised Main Text (Line 306-310):

To further assess robustness, we evaluated inter-batch reproducibility and examined the impact of pre-extraction room-temperature holding time. Both intra-plate and inter-batch variations remained low, and the decoded scores showed no appreciable drift within the tested handling-time window (**Supplementary Figure 10**).

Revised Main Text (Line 653-657):

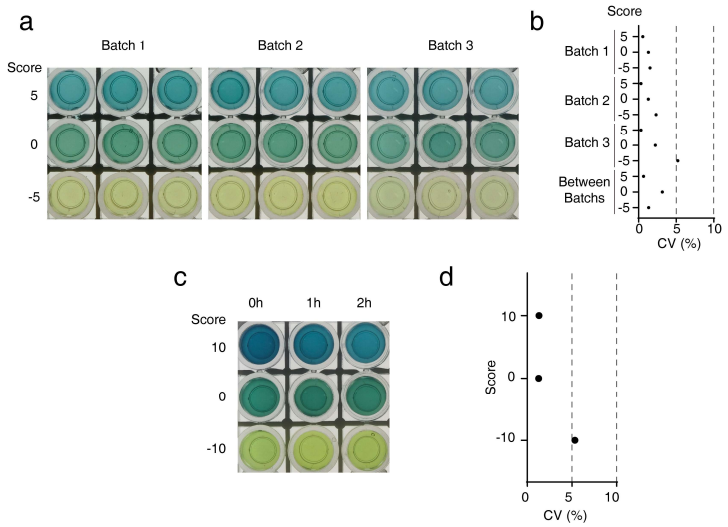
All assays were conducted under controlled ambient conditions (25 ± 1 °C, 50 ± 5% relative humidity). For image acquisition, a fixed-distance holder was used to standardize the camera-to-plate geometry, together with uniform diffuse illumination (LED ring light or backlight lightbox) to reduce shadows and spatial inhomogeneity. Each run included standardized reference wells for plate-to-plate calibration.

Revised Main Text (Line 741-745):

For calibration and robustness evaluation, a panel of score-referenced standard samples spanning the intended score range was processed in parallel with each assay plate. These standards

were measured in triplicate and served as internal controls to correct for run-to-run variations in enzyme activity and optical readouts, enabling consistent transformation of colorimetric signals into decoded scores.

Revised Supplementary Figure 10:



Supplementary Figure 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 (d) across different holding times and risk-score levels.

2. The rationale behind selecting the five miRNA markers is insufficient. The manuscript would benefit from a detailed description of the selection pipeline, supporting evidence from the literature, and a discussion of their biological roles in pancreatic cancer development.

Response:

We thank you for requesting clarification on the rationale for selecting the five miRNA markers. Briefly, we first performed differential expression screening in two public cohorts (GEO datasets GSE211692 and GSE163031) using the limma framework with predefined thresholds. We then retained miRNAs showing consistent dysregulation directionality across serum and tissue datasets to improve biological plausibility and reduce dataset-specific bias. Finally, we applied forward stepwise selection in a binary logistic regression model to derive a compact diagnostic signature, resulting in the following score function: Score = 0.43×miR-154-5p + 0.57×miR-629-5p + 0.41×miR-99a-5p – (0.09×miR-5006-5p + 0.51×miR-575).

To strengthen the biological and literature support, we have also added a brief discussion of the roles of these miRNAs in pancreatic cancer. Prior studies have reported miR-629 upregulation in pancreatic cancer and its contribution to proliferation and invasion, potentially through negative regulation of FOXO3. miR-99a-5p has been implicated in PDAC progression through the NOTCH2/Notch signaling pathway and has been proposed as a candidate prognostic/diagnostic

biomarker in multiple cohorts. Although the miR-154 family is frequently reported as tumor-suppressive in several malignancies, dysregulation can be context-dependent across tumor types; importantly, our analyses consistently identified miR-154-5p as a discriminative feature in pancreatic cancer cohorts, leading to its retention by the regression model. miR-575 has also been reported as a potential biomarker for pancreatic cancer detection. Finally, miR-5006-5p is relatively less studied in cancer, but emerging evidence suggests its involvement in tumor-related processes (e.g., via lncRNA-associated mechanisms). Collectively, these additions clarify both the data-driven selection logic and the biological plausibility of the final five-miRNA panel.

We do not claim this logistic-regression signature to be the globally optimal classifier; rather, we selected a transparent, literature- and dataset-supported model as a methodological benchmark to validate EnCODE's multi-miRNA encoding/decoding and risk-score computation. (**Line 445-453 and References 37-45**)

Revised Main Text (Line 445-453):

Prior studies have reported miR-629 upregulation in pancreatic cancer and its contribution to proliferation and invasion, potentially through negative regulation of FOXO3[37]. miR-99a-5p has been implicated in PDAC progression through the NOTCH2/Notch signaling pathway and has been proposed as a candidate prognostic/diagnostic biomarker in multiple cohorts[38-40]. Although the miR-154 family is frequently reported as tumor-suppressive in several malignancies, dysregulation can be context-dependent across tumor types[41, 42]. miR-575 has also been reported as a potential biomarker for pancreatic cancer detection[43,44]. Finally, miR-5006-5p is relatively less studied in cancer, but emerging evidence suggests its involvement in tumor-related processes (e.g., via lncRNA-associated mechanisms)[45].

Revised References 37-45:

37. Yan, H. et al. MiR-629 promotes human pancreatic cancer progression by targeting FOXO3. *Cell Death Dis* **8**, e3154 (2017).
38. Xu, J. et al. LncRNA MIR99AHG mediated by FOXA1 modulates NOTCH2/Notch signaling pathway to accelerate pancreatic cancer through sponging miR-3129-5p and recruiting ELAVL1. *Cancer Cell Int* **21**, 674 (2021).
39. Prinz, C., Fehring, L. & Frese, R. MicroRNAs as Indicators of Malignancy in Pancreatic Ductal Adenocarcinoma (PDAC) and Cystic Pancreatic Lesions. Cells 11 (2022).
40. Vietsch, E.E. et al. Immune-Related Circulating miR-125b-5p and miR-99a-5p Reveal a High Recurrence Risk Group of Pancreatic Cancer Patients after Tumor Resection. *Appl Sci (Basel)* **9** (2019).
41. Lin, C. et al. Oncogene miR-154-5p regulates cellular function and acts as a molecular marker with poor prognosis in renal cell carcinoma. *Life Sci* **209**, 481-489 (2018).
42. Cazzoli, R. et al. microRNAs derived from circulating exosomes as noninvasive biomarkers for screening and diagnosing lung cancer. *J Thorac Oncol* **8**, 1156-1162 (2013).
43. Sebastian, N.T. et al. Development of a MicroRNA Signature Predictive of Recurrence and Survival in Pancreatic Ductal Adenocarcinoma. *Cancers (Basel)* **13** (2021).
44. Imamura, T. et al. Urinary microRNA-210-3p as a novel and non-invasive biomarker for the detection of pancreatic cancer, including intraductal papillary mucinous carcinoma. *BMC Cancer* **24**, 907 (2024).

45. Wang, R., Wang, X., Zhang, J. & Liu, Y. LINC00942 Promotes Tumor Proliferation and Metastasis in Lung Adenocarcinoma via FZD1 Upregulation. *Technol Cancer Res Treat* **20**, 1533033820977526 (2021).

3. The process for determining the coefficients in the weighted calculation formula is unclear. The authors should describe the optimization strategy in detail (e.g., machine learning algorithms, statistical modeling), and provide justification for the chosen approach.

Response:

We thank you for this insightful comment. In the revised manuscript, we have added a detailed description of the coefficient estimation and optimization framework.

To this end, the coefficients were determined using binary logistic regression to model the probability of pancreatic cancer as a function of the selected miRNA features. Specifically, we log10-transformed miRNA concentrations to improve numerical stability and match the dynamic characteristics of the pre-amplification. We adopted a forward stepwise feature selection strategy to derive a parsimonious model, and the final coefficients were chosen to balance model discrimination and robustness while preserving interpretability. The full modeling procedure has now been clarified in the revised Methods section. (**Line 780-789**)

Revised Main Text (Line 780-789):

Binary logistic regression was used to construct a probability model for pancreatic cancer using the selected miRNAs as predictors. Prior to model fitting, miRNA concentrations were log10-transformed to reduce skewness and improve numerical stability across the wide dynamic range introduced by the pre-amplification step. Regression coefficients were estimated by maximum likelihood using the standard iterative algorithm implemented in SPSS (Version 20). A parsimonious model was derived using forward stepwise selection in SPSS (Forward: Wald), with candidate variables sequentially entered according to prespecified significance criteria (*entry* $p < 0.05$). Model discrimination was evaluated by receiver operating characteristic (ROC) analysis and summarized by the area under the curve (AUC). The final score was reported as a linear combination of the retained miRNAs with their estimated coefficients and directions of effect.

4. A direct comparison with existing point-of-care testing (POCT) platforms is missing. To highlight the novelty and translational value of EnCODE, please provide comparative data with at least two commercially available rapid detection methods, including cost-effectiveness and operational complexity assessments.

Response:

We thank you for the constructive suggestion. To address this point and better contextualize the translational value of EnCODE, we have added a direct comparison with three representative commercially available rapid NAAT systems: Roche cobas® Liat (cartridge-based rapid PCR), Cepheid GeneXpert Xpress® (rapid cartridge-based RT-PCR), and Abbott ID NOW™ (isothermal NAAT based on NEAR technology). These platforms collectively cover the major categories of widely deployed rapid molecular POCT in clinical practice.

Overall, cartridge-based NAAT/RT-PCR POCT systems (e.g., Roche cobas® Liat, Cepheid

GeneXpert Xpress®, and Abbott ID NOW™) offer the shortest sample-to-answer time (typically ~5–30 min) and highly streamlined "sample-in, result-out" workflows. However, these platforms are commonly optimized for single-sample or low-throughput testing, provide fixed and proprietary assay menus, and can incur higher per-test consumable costs, especially when multi-target panels are desired. In contrast, EnCODE introduces a distinct colorimetric encoding–decoding paradigm that integrates multi-miRNA information within a single reaction, directly compressing multiplex molecular signals into interpretable optical codes and quantitative risk scores. Because information is encoded in a controllable color space rather than multiple fluorescent channels, EnCODE enables multi-target integration without being constrained by spectral overlap in conventional fluorescence multiplexing, and it supports plate-based parallelization (32–96 samples per run) that aligns with standard laboratory instrumentation and scalable automation (e.g., generic plate handlers and liquid-handling workflows). In settings where instrumentation is limited, EnCODE readout can also be simplified to smartphone imaging for rapid, low-cost semi-quantitative interpretation. Collectively, these characteristics make EnCODE a novel approach for high-density multi-marker molecular profiling and cost-effective batch testing, positioning it as a distinct solution compared to conventional POCT platforms. **(Line 578-582)**

Revised Main Text (Line 578-582):

Compared with commercial rapid NAAT/RT-PCR POCT platforms (e.g., Roche cobas® Liat, Cepheid GeneXpert Xpress®, and Abbott ID NOW™), EnCODE offers a unique colorimetric-based technology with a focus on multiplex information density, flexible marker integration, and plate-based batch throughput. This technology is particularly advantageous for batch screening and multi-marker risk scoring scenarios.

5. The current study includes only 32 samples (12 pancreatic cancer patients), which is insufficient to draw reliable conclusions. The authors are strongly encouraged to increase the sample size to a statistically meaningful scale and provide a clear rationale with sample size calculations.

Response:

We thank you for this important comment. To address this concern, we performed a binomial-proportion–based sample size estimation using clinically meaningful target metrics and have expanded the cohort accordingly.

Based on our pilot data and prior literature, we set the target sensitivity (Se) at 90% and specificity (Sp) at 85%. The required numbers of disease-positive (N_pos) and disease-negative (N_neg) samples were approximated using the standard normal approximation for estimating a proportion with a desired precision:

$$N = \frac{Z_{1-\frac{\alpha}{2}}^2 \times p(1-p)}{d^2}$$

where $Z_{1-\alpha/2}$=1.96 for a two-sided 95% confidence interval, p represents the expected Se or Sp, and d is the allowable absolute error (half-width of the CI). Using d=7.5%, the estimated minimum sample size is approximately 62 pancreatic cancer cases (for Se = 0.90) and 87 controls (for Sp = 0.85), resulting in a total cohort size of ~149 samples to achieve ±7.5% precision for both sensitivity and specificity.

We assembled an expanded discovery cohort comprising 63 pancreatic cancer cases, 50 healthy controls, and 50 disease controls (pancreatitis), which meets the calculated requirement under the ±7.5% error criterion. The revised manuscript and Supplementary Information now clearly describe the enlarged cohort and EnCODE performance on the discovery cohort (**Line 454-463, Figure 5 and Supplementary Figure 16**).

Revised Main Text (Line 454-463):

Finally, we evaluated EnCODE using a clinical serum cohort comprising pancreatic cancer (n = 63), benign pancreatic disease controls (pancreatitis, n = 50), and healthy controls (n = 50) (**Figure 5f, Supplementary Table 4**). EnCODE generated directly interpretable prediction scores and enabled robust classification of pancreatic cancer versus controls, achieving an AUC (area under the ROC curve) of 0.9586 (95% CI: 0.9310–0.9862), with 96.83% sensitivity (95% CI: 89.14–99.44%) and 86.00% specificity (95% CI: 77.86–91.47%). Importantly, EnCODE provide effective discrimination from benign pancreatic disease (pancreatitis) (**Figure 5g–j**). In parallel, qRT-PCR profiling of the same miRNA panel recapitulated the expression differences in clinical samples and showed comparable discrimination performance, supporting the analytical consistency between EnCODE and conventional PCR-based workflows (**Supplementary Figure 16b–d**).

Revised Figure 5:

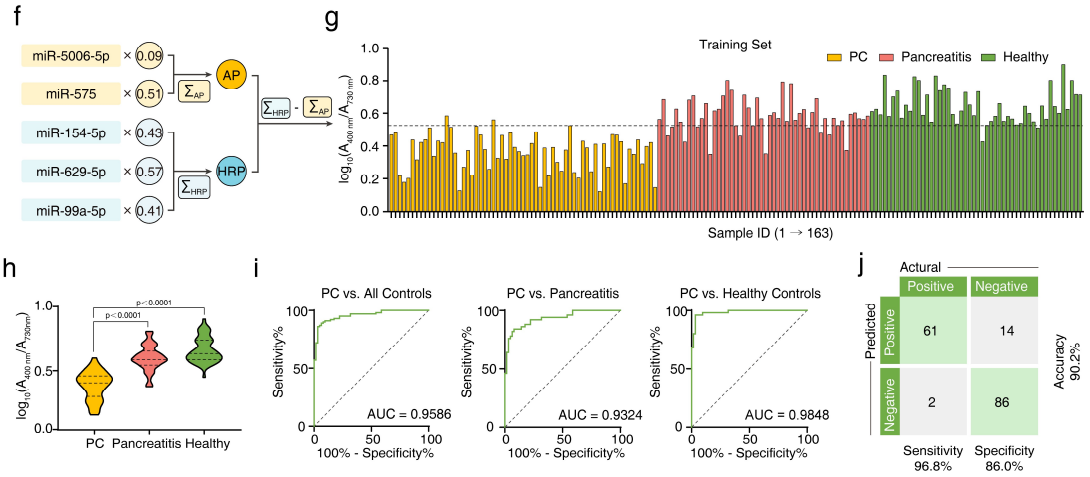
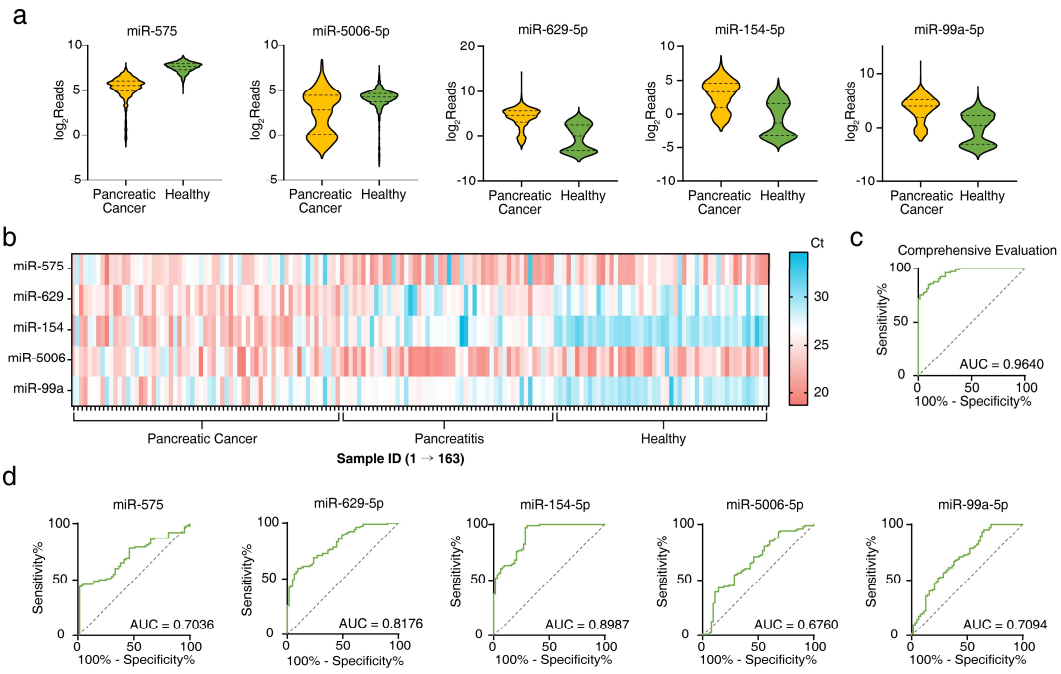


Figure 5. (f) Schematic diagram illustrates the pancreatic cancer prediction model: Score = 0.43×miR-154-5p + 0.57×miR-629-5p + 0.41×miR-99a-5p – (0.09×miR-5006-5p + 0.51×miR-575). (g) Representative EnCODE readouts for 163 clinical serum samples in the expanded cohort, including 63 pancreatic cancer cases, 50 pancreatitis controls, and 50 healthy controls. (h) Risk-score–based prediction results for pancreatic cancer classification derived from EnCODE measurements in the expanded clinical cohort. (i) ROC curve of EnCODE performance for clinical sample classification. (j) Confusion matrix summarizing classification accuracy of EnCODE in the expanded clinical cohort.

Revised Supplementary Figure 16:



Supplementary Figure 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.

6. An independent external validation cohort is missing. To establish the robustness and generalizability of the EnCODE platform, data from additional medical centers and different populations should be included.

Response:

Thank you for this important point regarding the need for independent external validation. In response, we have added an independent external validation cohort to evaluate the robustness and generalizability of the EnCODE platform. Specifically, in addition to the discovery/training cohort collected at Shanghai Tenth People's Hospital, we established a separate validation cohort collected from two additional medical centers (Shanghai Fourth People's Hospital and Wuxi People's Hospital), thereby enabling cross-center assessment in distinct patient populations.

The results from this external validation cohort are now presented in the revised manuscript. To further assess generalizability, we tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15). EnCODE showed consistent performance between the discovery and external validation cohorts, supporting its reproducibility and generalizability across medical centers. (**Line 463-466, Line 791-792 and Figure 5**).

Revised Main Text (Line 463-466):

To further assess generalizability, we tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15), which yielded consistent classification accuracy (**Supplementary Table 5, Figure 5k, l**).

Revised Main Text (Line 791-792):

All serum samples were collected at the 10th People's Hospital of Shanghai (Shanghai, China), Shanghai Fourth People's Hospital (Shanghai, China) and Wuxi People's Hospital (Wuxi, China).

Revised Figure 5:

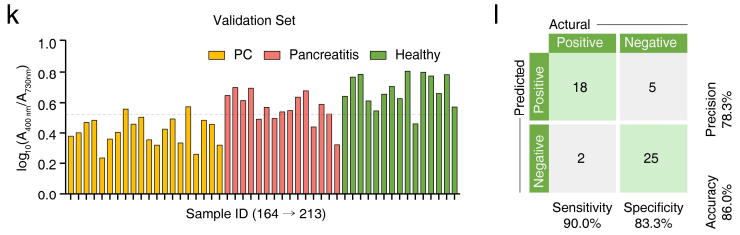


Figure 5. (k) Representative EnCODE readouts for an independent external validation cohort of 50 clinical serum samples collected from additional medical centers, including 20 pancreatic cancer cases, 15 pancreatitis controls, and 15 healthy controls. (l) Confusion matrix summarizing EnCODE classification accuracy in the external validation cohort. Data are presented as the mean ± s.d. from three independent experiments (n = 3).

7. The control group design is suboptimal, as it does not incorporate benign pancreatic diseases (e.g., chronic pancreatitis). Including such cases is essential for accurately evaluating diagnostic specificity.

Response:

Thank you for this important suggestion regarding control-group design. In the revised study, we expanded the control arm to include a dedicated benign pancreatic disease cohort (pancreatitis) as disease controls, enabling a more clinically relevant evaluation of diagnostic specificity and differential diagnostic utility. The updated analyses now report EnCODE performance in distinguishing pancreatic cancer not only from healthy controls but also from benign pancreatic diseases. Specifically, we assembled an expanded discovery cohort comprising 63 pancreatic cancer cases, 50 healthy controls, and 50 pancreatitis controls. Further, we additionally tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15). (**Line 454–466 and Figure 5**).

Revised Main Text (Line 454-466):

Finally, we evaluated EnCODE using a clinical serum cohort comprising pancreatic cancer (n = 63), benign pancreatic disease controls (pancreatitis, n = 50), and healthy controls (n = 50) (**Figure 5f, Supplementary Table 4**). EnCODE generated directly interpretable prediction scores and enabled robust classification of pancreatic cancer versus controls, achieving an AUC (area under the ROC curve) of 0.9586 (95% CI: 0.9310–0.9862), with 96.83% sensitivity (95% CI: 89.14–99.44%) and 86.00% specificity (95% CI: 77.86–91.47%). Importantly, EnCODE provide effective discrimination from benign pancreatic disease (pancreatitis) (**Figure 5g–j**). In parallel, qRT-PCR

profiling of the same miRNA panel recapitulated the expression differences in clinical samples and showed comparable discrimination performance, supporting the analytical consistency between EnCODE and conventional PCR-based workflows (**Supplementary Figure 16b–d**). To further assess generalizability, we tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15), which yielded consistent classification accuracy (**Supplementary Table 5, Figure 5k, l**).

Revised Figure 5:

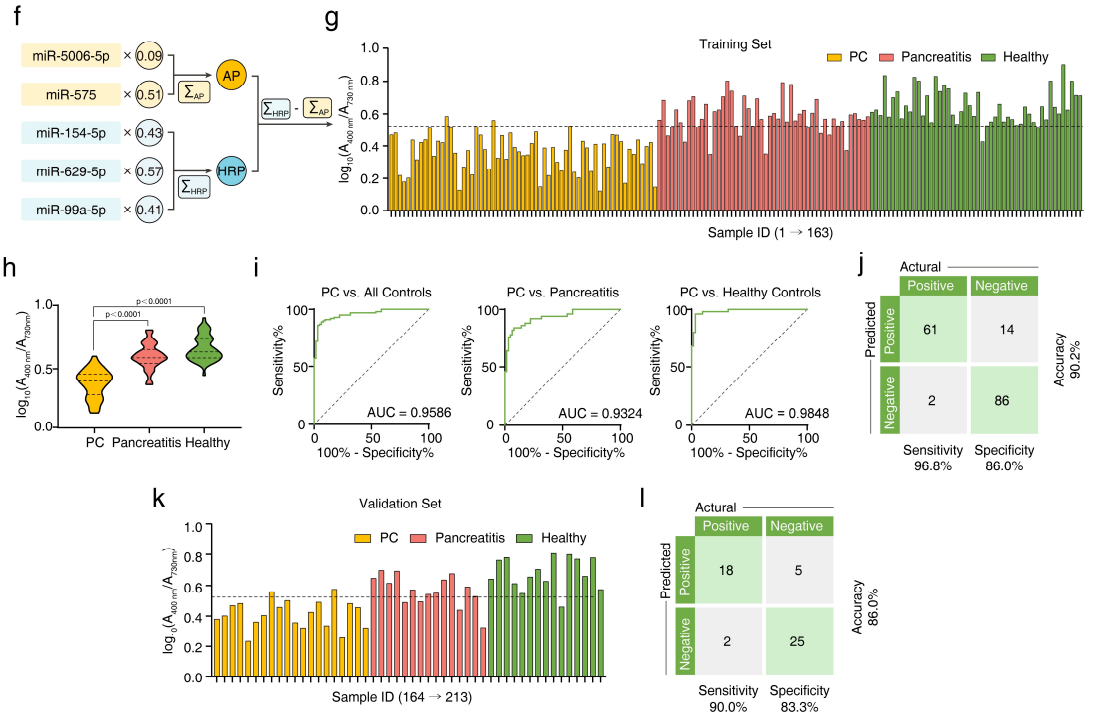


Figure 5. (f) Schematic diagram illustrates the pancreatic cancer prediction model: Score = 0.43×miR-154-5p + 0.57×miR-629-5p + 0.41×miR-99a-5p – (0.09×miR-5006-5p + 0.51×miR-575). (g) Representative EnCODE readouts for 163 clinical serum samples in the expanded cohort, including 63 pancreatic cancer cases, 50 pancreatitis controls, and 50 healthy controls. (h) isk-score–based prediction results for pancreatic cancer classification derived from EnCODE measurements in the expanded clinical cohort. (i) ROC curve of EnCODE performance for clinical sample classification. (j) Confusion matrix summarizing classification accuracy of EnCODE in the expanded clinical cohort. (k) Representative EnCODE readouts for an independent external validation cohort of 50 clinical serum samples collected from additional medical centers, including 20 pancreatic cancer cases, 15 pancreatitis controls, and 15 healthy controls. (l) Confusion matrix summarizing EnCODE classification accuracy in the external validation cohort. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Statistical analysis was performed using a Student's t-test.

8. Critical clinical information (e.g., disease stage, pathological type) is lacking for the enrolled patients. Please provide complete baseline clinical characteristics.

Response:

Thank you for this valuable comment. In the revised manuscript, we have included the complete baseline clinical characteristics for all enrolled patients, covering essential clinical information such as disease stage, pathological type, tumor location, age, sex, and relevant comorbidities. (**Line 454-456, Supplementary Table 4 and Supplementary Table 5**)

Revised Main Text (Line 454-456):

Finally, we evaluated EnCODE using a clinical serum cohort comprising pancreatic cancer (n = 63), benign pancreatic disease controls (pancreatitis, n = 50), and healthy controls (n = 50) (**Figure 5f, Supplementary Table 4**).

Revised Supplementary Table 4:

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	/	/

Revised Supplementary Table 5:

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: 7 50-60: 7 40-50: 1 30-40: 2	70-80: 1 60-70: 10 50-60: 6 40-50: 3
Sex	Male: 9 Female: 11	Male: 10 Female: 10	Male: 10 Female:10
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	/	/

9. The GSE163031 dataset used for validation contains too few samples, raising concerns about potential overfitting. A larger external dataset should be used for validation, or appropriate overfitting tests should be presented.

Response:

We thank you for this thoughtful comment regarding the limited sample size of GSE163031 and the resulting concern about potential overfitting. In our study GSE163031 was not used as the sole validation dataset, but rather as one of multiple independent external cohorts incorporated to assess robustness.

Specifically, GSE211692 (serum cohort, n=16190) and GSE163031 (tissue cohort, n =15) were used for biomarker screening and model construction, where we retained miRNAs showing consistent dysregulation directionality across serum and tissue to enhance biological plausibility. Importantly, the final miRNA panel and prediction model were further evaluated on an independent external validation cohort, GSE106817(n=4046), which was processed and analyzed separately from the training procedure to avoid any information leakage between training and validation steps. The consistent performance observed in this independent dataset supports the robustness and generalizability of the proposed miRNA marker panel. (**Line 431-444 and Figure 5**)

Revised Main Text (Line 431-444):

After determining the clinical testing workflow, we applied EnCODE to pancreatic cancer diagnosis and prediction. We first attempted to construct a pancreatic cancer diagnosis prediction model. Using the GSE211692 dataset (included serum miRNA expression levels from healthy controls and pancreatic cancer patients) and the GSE163031 dataset (included pancreatic tissue miRNA expression levels from normal and cancerous samples) from the Gene Expression Omnibus (GEO) database, we identified three upregulated and two downregulated miRNAs consistently expressed across both sample types (**Figure 5c, Supplementary Figure 16a**). Subsequently, a

pancreatic cancer diagnosis model was constructed via binary logistic regression: Score = 0.43 × miR-154-5p + 0.57 × miR-629-5p + 0.41 × miR-99a-5p − (0.09 × miR-5006-5p + 0.51 × miR-575). Notably, initial miRNA concentrations were log10-transformed to make Notably, initial miRNA concentrations were log10-transformed to make the concentration data compatible with the LATE-PCR amplification mechanism. The model demonstrated strong discrimination between pancreatic cancer and healthy samples in public datasets (**Figure 5d**) and performed robustly across serum (GSE106817, GSE211692) and tissue (GSE163031) cohorts (**Figure 5e**).

Revised Figure 5:

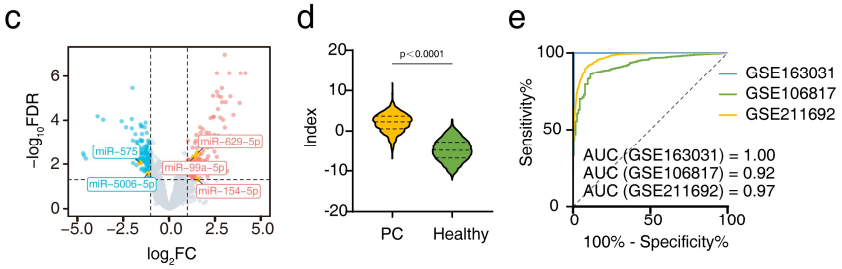


Figure 5. (c) Volcano plot of differentially expressed miRNAs from the GSE163031 dataset. Yellow dots indicate miRNAs selected for the pancreatic cancer prediction model. (d) Plot of pancreatic cancer prediction scores (GSE211692 dataset) for healthy vs. patient groups using the GSE211692 dataset. (e) The ROC curve validating the pancreatic cancer diagnosis model, with data from the GSE106817, GSE211692, and GSE163031 datasets.

Minor Revisions

10. The manuscript should include a detailed description of how the concentrations of EnCODE components (primers, templates, enzymes) were optimized and determined.

Response:

We thank you for this comment. In the revised manuscript, we now provide a detailed description of how the working concentrations of key EnCODE components (capture oligonucleotide, RCA template, and enzyme–DNA conjugates) were optimized and determined.

The capture oligonucleotide was immobilized on a streptavidin-coated microplate surface. To maximize surface coverage and minimize well-to-well variability, we adopted the manufacturer-recommended saturating coating concentration, which reliably saturates the available streptavidin binding sites and was therefore used as the final capture-probe concentration in EnCODE.

The working concentration of the template strand was selected to satisfy two practical criteria: (i) the template concentration should be at least one order of magnitude higher than the post-amplification target level to ensure efficient circularization and downstream RCA; and (ii) under equilibrium binding, the immobilized capture probes should be able to fully capture the template molecules to avoid excess free template. We evaluated a range of template concentrations and selected 50 nM, which met both criteria and produced stable RCA yields with consistent downstream decoding performance.

To determine the optimal concentration of the enzyme–DNA conjugate, we performed EnCODE measurements at a fixed target level while titrating the conjugate over a broad range (0–100 nM). Absorbance increased proportionally at low-to-intermediate concentrations and

approached saturation above ~25 nM. To ensure near-saturating binding/signal generation and thereby maximize robustness and minimize sensitivity to small concentration fluctuations, we selected 100 nM as the working concentration (**Supplementary Figure 9**). These optimization details have now been added to the revised manuscript and Supplementary Information. (**Line 254-260, Line 714-737 and Supplementary Figure 9**).

Revised Main Text (Line 254-260):

We optimized the concentration of key components in the EnCODE assay. The capture oligonucleotide was immobilized at the manufacturer-recommended saturating concentration to ensure maximal surface coverage and minimize variability. The template strand concentration was selected at 50 nM, which met two criteria: ensuring efficient circularization while enabling complete capture by the immobilized probes. For the enzyme–DNA conjugates, we titrated 0–100 nM and selected 100 nM to ensure near-saturating signal generation and robust readout. (**Supplementary Figure 9**).

Revised Main Text (Line 714-737):

Application of Enzymatic Colorimetric Coding Detection Process for Target miRNA

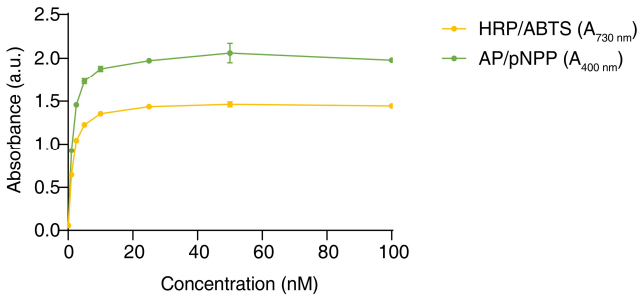
Firstly, the circularization of the template DNA is achieved by preparing a 100 μL reaction mixture that includes 1× SplintR Ligase buffer, miRNA samples, 50 nM of the corresponding template DNA, and 40 U/μL of SplintR Ligase. This mixture is incubated at 25°C for 30 minutes to form the circular template DNA-target heterodimer.

Concurrently, the capture strand is immobilized on a streptavidin-coated 96-well enzyme plate by adding 10 μM of 5'-biotinylated capture strand to each well and incubating at room temperature for 2 hours. The wells are washed three times with PBS to eliminate excess capture strand, ensuring complete immobilization.

The circular template DNA-target complex is then incubated at room temperature for 30 minutes within the wells coated with the capture strand, followed by three washes. Subsequently, rolling circle amplification (RCA) is performed by mixing 200 μL of 0.4 mM dNTP, 0.2 U/μL phi29 DNA polymerase, 100 nM HRP-DNA barcode 1, and 100 nM AP-DNA barcode 2 in a 200 μL 1× phi29 DNA polymerase buffer system within the microplate wells and incubating at 30°C for 30 minutes.

Lastly, an enzyme-catalyzed colorimetric reaction is conducted by adding 200 μL of the color development solution, which consists of 5 mM pNPP, 5 mM ABTS, and 0.5 mM H_2O_2 mixed in 1× PBS, to the enzyme plate. The reaction is allowed to proceed at room temperature for 30 minutes. The absorbance measurements are taken using a multi-mode microplate reader (Infinite 200 PRO, Tecan, Switzerland), and the color development results are photographed and analyzed. All assays were conducted under controlled ambient conditions (25 ± 1 °C, 50 ± 5% relative humidity). For image acquisition, a fixed-distance holder was used to standardize the camera-to-plate geometry, together with uniform diffuse illumination (LED ring light or backlight lightbox) to reduce shadows and spatial inhomogeneity.

Revised Supplementary Figure 9:



Supplementary Figure 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).

11. Please clarify the optimization strategy for the coefficients in the weighted calculation formula, and indicate whether alternative mathematical models were tested and how their performance compared.

Response:

We thank you for this insightful comment. In the revised manuscript, we have added a detailed description of the coefficient estimation.

We emphasize that the present work is primarily a methodological study, aimed at establishing a robust and metrologically stable signal–decoding workflow for the EnCODE platform, rather than developing a data-hungry clinical prediction algorithm. Accordingly, our objective in coefficient determination was to maximize decoding integrity, linearity, and robustness, while maintaining full interpretability.

To this end, the coefficients were determined using binary logistic regression to model the probability of pancreatic cancer as a function of the selected miRNA features. Specifically, we log10-transformed miRNA concentrations to improve numerical stability and match the dynamic characteristics of the pre-amplification. We adopted a forward stepwise feature selection strategy to derive a parsimonious model, and the final coefficients were chosen to balance model discrimination and robustness while preserving interpretability. The full modeling procedure has now been clarified in the revised Methods section. (**Line 780-789**)

Revised Main Text (Line 780-789):

Binary logistic regression was used to construct a probability model for pancreatic cancer using the selected miRNAs as predictors. Prior to model fitting, miRNA concentrations were log10-transformed to reduce skewness and improve numerical stability across the wide dynamic range introduced by the pre-amplification step. Regression coefficients were estimated by maximum likelihood using the standard iterative algorithm implemented in SPSS (Version 20). A parsimonious model was derived using forward stepwise selection in SPSS (Forward: Wald), with candidate variables sequentially entered according to prespecified significance criteria (*entry* $p < 0.05$). Model discrimination was evaluated by receiver operating characteristic (ROC) analysis and summarized by the area under the curve (AUC). The final score was reported as a linear combination of the retained miRNAs with their estimated coefficients and directions of effect.

12. Additional validation data demonstrating the linear relationship between color signals and miRNA concentrations, including the linear detection range, should be provided.

Response:

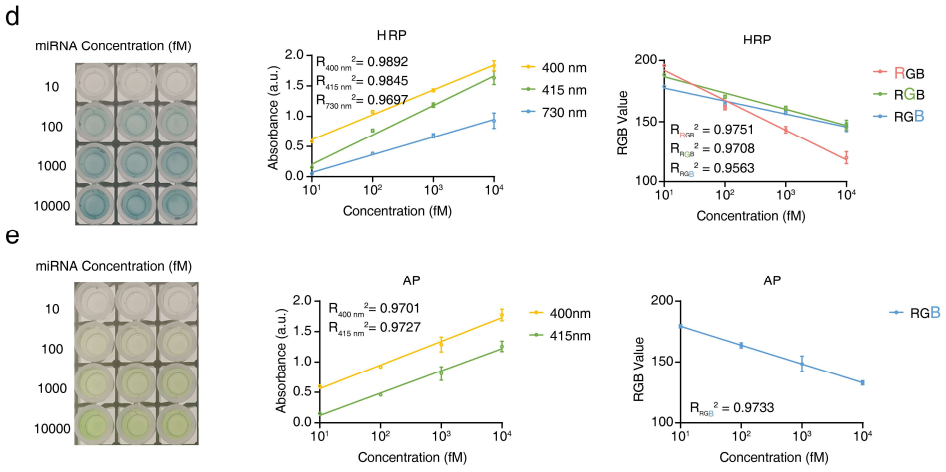
We thankyou for this valuable comment. In the revised manuscript, we have included comprehensive validation data demonstrating the linear relationship between miRNA concentrations and EnCODE colorimetric signals.

We measured the target miRNA samples using EnCODE. Consistent with the clinical-sample workflow, the target miRNAs underwent the same pre-amplification step prior to EnCODE readout. Across the tested range (10 fM to 10 pM), both the spectral readout and the colorimetric (RGB-based) readout showed strong linear relationships with the logarithm of miRNA abundance, with $R^2 > 0.98$ for the spectral features and $R^2 > 0.97$ for the color features. These results support a robust linear mapping between EnCODE optical signals and miRNA concentration over a 10 fM–10 pM dynamic range (**Supplementary Figure 13**). (**Line 406-413 and Supplementary Figure 13**)

Revised Main Text (Line 406-413):

To further assess metrological stability in a workflow-relevant setting, we benchmarked EnCODE using miRNA samples processed through the same pre-amplification step as in the clinical workflow. Across the tested concentration series, EnCODE readouts closely followed the input miRNA levels and showed strong agreement with qRT-PCR quantification, supporting reliable quantitative transfer from miRNA abundance to enzymatic output (**Supplementary Figure 13a–c**). In parallel, both the HRP–ABTS and AP–pNPP systems exhibited robust linear relationships between miRNA concentration and colorimetric features at the levels of absorbance and RGB values, further supporting a stable miRNA-to-enzyme signal mapping (**Supplementary Figure 13d-e**).

Revised Supplementary Figure 13:



Supplementary Figure 13. (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).

13. The effects of sample storage conditions (e.g., temperature, storage time) on assay performance should be systematically assessed.

Response:

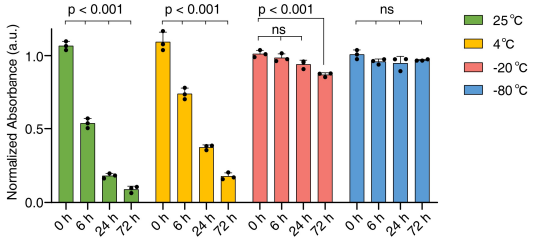
We thank you for highlighting the importance of sample stability. In the revised manuscript, we have systematically assessed the effects of storage temperature and storage time on EnCODE assay performance.

To systematically assess the effects of sample storage temperature and duration, we used serum derived from a single donor specimen and aliquoted it to minimize variability arising from endogenous miRNA differences across samples. The aliquots were stored under common conditions (4 °C, −20 °C, and −80 °C) and analyzed at predefined time points (0 h, 6 h, 24 h, and 72 h) using the complete EnCODE workflow. Based on these measurements, we defined practical storage limits for reliable quantification: storage at −20 °C for up to 24 h or at −80 °C for up to 72 h resulted in signal deviations within ±10% relative to freshly processed controls. These stability constraints have now been incorporated into the revised manuscript as actionable guidelines for sample handling and pre-analytical quality control. This systematic evaluation shows that EnCODE remains robust under typical storage conditions, as long as pre-analysis handling is within the defined stability window (**Line 425-430 and Supplementary Figure 15**).

Revised Main Text (Line 425-430):

In addition, we systematically assessed the effects of storage conditions (4°C, -20°C, -80°C) and storage duration (0–72 hours) on the accuracy of miRNA detection. Our results showed that miRNA signals remained stable when samples were stored at -20°C for up to 1 day or at -80°C for up to 3 days, with no significant deviation from fresh controls. These findings provide clear guidance on sample storage conditions for reliable EnCODE analysis (**Supplementary Figure 15**).

Revised Supplementary Figure 15:



Supplementary Figure 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 Student's t-test, and p values are indicated.

14. Please provide a quantitative analysis of how common interfering factors (e.g., hemolysis, lipemia) affect assay accuracy.

Response:

We sincerely thank you for raising the important issue. We have conducted targeted interference experiments to evaluate these effects and developed actionable rules for sample rejection and correction based on our findings.

Although the EnCODE workflow primarily focuses on the extraction of free miRNA from plasma/serum samples, which does not directly affect HRP/AP reactions or optical readouts, sample conditions may impact miRNA extraction efficiency and, in turn, affect the final results.

To assess potential interference, we established three controlled experimental groups using blood samples from the same source, thereby minimizing variability arising from endogenous miRNA differences across individuals. First, to evaluate anticoagulants and coagulants, blood was collected into commonly used collection tubes (e.g., heparin- or EDTA-containing tubes, and serum tubes with clot activator), and EnCODE readouts were compared across tube conditions. Second, to assess interference from pancreatic cancer therapeutics, matched aliquots of the same blood sample were supplemented with representative drugs (5-Fluorouridine, Irinotecan, Oxaliplatin, and Gemcitabine) at the tested concentrations prior to processing. Third, to model common matrix interferences, equal volumes of Intralipid, albumin solution, or red-blood-cell lysate were added to matched aliquots to simulate lipemia, hyperproteinemia, and hemolysis, respectively.

All groups were processed through the complete EnCODE workflow, and outputs were compared against the corresponding untreated aliquot controls (**Supplementary Figure 14a**). Anticoagulants/coagulants and the tested therapeutic agents did not produce appreciable changes in EnCODE readouts, and neither lipemic nor high-protein conditions showed significant interference. In contrast, hemolysis led to a modest upward shift in measured miRNA levels, consistent with the release of endogenous miRNAs from lysed erythrocytes.

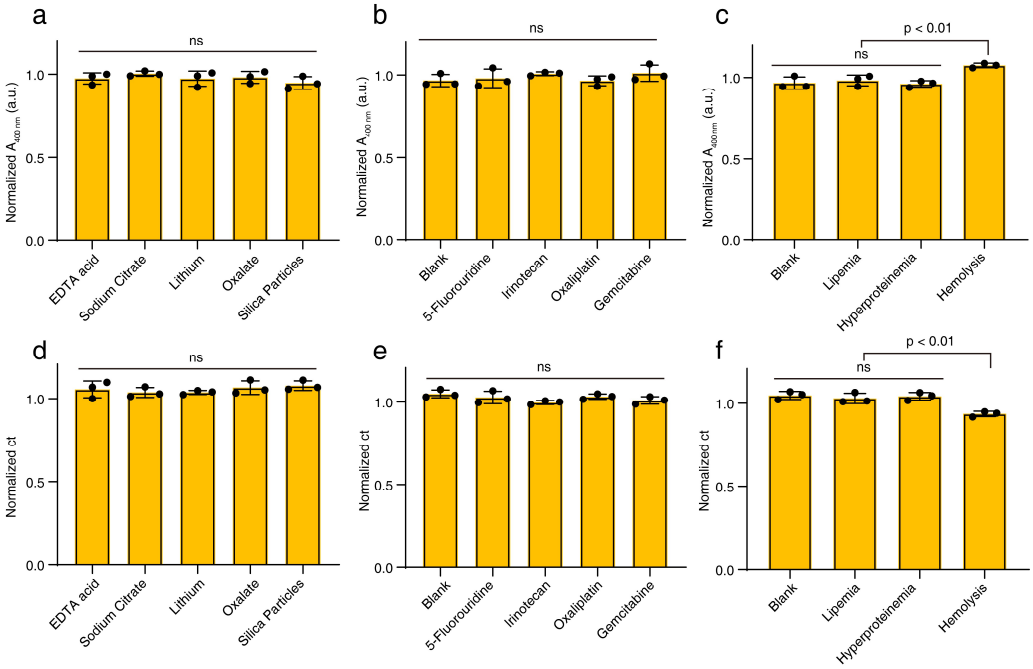
In addition, EnCODE readouts showed the same directional trends as those obtained by qRT-PCR on the matched samples, supporting that the tested sample conditions do not materially bias EnCODE-based quantification of miRNA abundance (**Supplementary Figure 14b**).

Based on these findings, we conclude that hemolytic samples can affect miRNA detection, and thus, careful evaluation of sample hemolysis should be conducted before processing. (**Line 414-423 and Supplementary Figure 14)**

Revised Main Text (Line 414-423):

To evaluate robustness under clinically relevant conditions, we assessed common pre-analytical interferences and therapeutic agents using aliquots from the same-source blood sample. We tested anticoagulants/coagulants, representative pancreatic cancer drugs (5-Fluorouridine, Irinotecan, Oxaliplatin, and Gemcitabine), and simulated matrix interferences (lipemia, high protein, and hemolysis) by adding equal volumes of Intralipid, albumin, or red-blood-cell lysate, respectively. Anticoagulants/coagulants, drugs, and lipemic/high-protein conditions showed no appreciable impact on EnCODE readouts, whereas hemolysis caused a modest upward shift consistent with erythrocyte miRNA release (**Supplementary Figure 14a-c**). Notably, EnCODE readouts followed the same directional trends as qRT-PCR on the matched samples, supporting that these sample conditions do not materially bias miRNA quantification (**Supplementary Figure 14d-f**).

Revised Supplementary Figure 14:



Supplementary Figure 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 Student's t-test, and p values are indicated.

15. Variability in subjective interpretation of colorimetric results should be evaluated, and strategies to improve consistency should be proposed.

Response:

We thank you for this valuable comment. In the revised manuscript, we have now systematically assessed inter-rater consistency and implemented objective procedures to improve readout robustness.

To minimize operator-dependent variation during image-based color readout, all sample handling and EnCODE procedures were performed under environmental chamber (25 ± 1 °C and 50 ± 5% relative humidity). For image-based readout, we additionally standardized the imaging geometry using a fixed-distance holder and employed uniform diffuse illumination (LED ring light or a backlight lightbox) to minimize shadows and spatial non-uniformity. Each assay run included internal reference standards for plate-to-plate normalization, thereby correcting potential variations in spectral response and enzyme activity. In practical deployment, we recommend reporting results primarily via objective RGB/spectral decoding (with visual scoring as an auxiliary output) to further

reduce subjectivity.

To quantitatively evaluate the variability of human visual interpretation, three blinded raters (n = 3) independently scored the same set of representative outputs spanning four intensity levels (strong positive, weak positive, strong negative, and weak negative). The interpretation results were highly consistent across the three raters, indicating that subjective judgment is unlikely to introduce major variability in result interpretation. These results, together with the standardized imaging and objective RGB/spectral decoding workflow, support that EnCODE readouts can be reliably interpreted with improved consistency. (**Line 302-306, Line 653-657 and Supplementary Table 1**)

Revised Main Text (Line 302-306):

To quantify the variability of subjective visual interpretation, three blinded raters independently classified representative EnCODE outputs spanning strong/weak positive and strong/weak negative cases, showing substantial inter-rater agreement (accuracy = 94.4%; **Supplementary Table 1**). This further supports the consistency of visual readout when complemented by standardized imaging and objective decoding.

Revised Main Text (Line 653-657):

All assays were conducted under controlled ambient conditions (25 ± 1 °C, 50 ± 5% relative humidity). For image acquisition, a fixed-distance holder was used to standardize the camera-to-plate geometry, together with uniform diffuse illumination (LED ring light or backlight lightbox) to reduce shadows and spatial inhomogeneity. Each run included standardized reference wells for plate-to-plate calibration.

Revised Supplementary Table 1:

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

16. Finally, please describe the pathway and technical feasibility of smartphone integration with the EnCODE platform.

Response:

We thank you for this valuable suggestion. In the revised manuscript, we have added a dedicated description of the pathway and technical feasibility for integrating EnCODE with smartphone-based readout to enhance field deployability.

Smartphone integration is naturally compatible with EnCODE because the final output is a stable, static colorimetric code, which does not require sophisticated optics or time-resolved measurements. Modern smartphone cameras (typically 8–12 MP or higher) can reliably capture well-level RGB information when illumination and imaging geometry are standardized, enabling quantitative color extraction with minimal deviation relative to benchtop imaging under controlled conditions (within ~5% in our tests). A practical prototype workflow requires only three low-cost components: (i) a fixed-distance holder to enforce consistent camera–plate geometry, (ii) a diffuse uniform lighting module (e.g., LED ring light or backlight lightbox) to suppress shadows and spatial non-uniformity, and (iii) a lightweight smartphone application that automatically segments wells, extracts RGB (or Hue) features, and maps them to calibrated EnCODE decoding coefficients.

To further improve robustness, the smartphone workflow can incorporate internal color references for plate-level normalization and automated quality-control checks (e.g., focus, exposure saturation, and illumination uniformity) prior to decoding. In operation, the user would only need to capture a single image per plate, and the readout, normalization, and real-time decoding can be completed within seconds (typically <30 s). Collectively, these readily implementable elements support the technical feasibility of smartphone-based EnCODE readout and provide a clear pathway toward portable, user-friendly deployment. **(Line 567-576)**

Revised Main Text (Line 567-576):

In particular, because EnCODE yields a stable, static colorimetric output, smartphone integration is technically straightforward and does not require specialized optics or time-resolved measurements. A practical workflow can be realized with standardized illumination and fixed imaging geometry (e.g., a simple holder plus diffuse LED lighting), allowing reliable well-level RGB extraction and on-device decoding via a lightweight app. Robustness can be further improved by incorporating internal color references for plate-level normalization and automated QC checks (focus, exposure, and illumination uniformity). In future iterations, computer-vision–assisted capture (auto-segmentation, perspective correction, and real-time guidance) could enable rapid, real-time readout and decoding within seconds.

Reviewer #3:

This manuscript presents EnCODE, an enzymatic colorimetric encoding system for detecting multiple miRNA biomarkers in clinical samples. While the core technical concept demonstrates some merit, substantial concerns limit the work's current contribution and reliability.

The authors report several noteworthy results, including the development of a dual-enzyme system that mathematically encodes biomarker concentrations into interpretable color patterns, demonstration of linear relationships between enzyme concentrations and spectral outputs, and extension to three-dimensional encoding. In a clinical study of 32 samples, they achieved reported sensitivity of 100% and overall accuracy of 91% for pancreatic cancer detection. However, these results require critical examination.

Regarding significance to the field, the work makes a meaningful contribution to colorimetric biosensing and multiplexed nucleic acid detection. The mathematical encoding framework represents a novel approach - while colorimetric detection exists widely, the systematic use of enzymatic color mixing as a mathematically rigorous information encoding system with reversible decoding capabilities is innovative. This differs substantially from conventional colorimetric approaches that typically provide simple positive/negative or semi-quantitative readouts. However, the clinical diagnostic performance does not clearly exceed established miRNA detection methods, and without comparative studies, the practical advantages over qRT-PCR or digital PCR remain unclear. The clinical application, while demonstrating feasibility, requires more robust validation to assess its true diagnostic value.

Response:

Thank you for the thoughtful evaluation and for recognizing the novelty of EnCODE as a mathematically defined enzymatic colorimetric encoding–decoding framework for multiplex nucleic-acid readout. In the revised manuscript, we have therefore (i) explicitly repositioned EnCODE as a methodological proof-of-concept platform rather than a ready-to-deploy diagnostic, (ii) expanded the experimental validation of the encoding/decoding framework and its operating limits, and (iii) strengthened the clinical evaluation by enlarging the cohort and adding an independent external validation set, with performance now reported alongside confidence intervals and appropriate caveats.

Based on our pilot data and prior literature, we set the target sensitivity (Se) at 90% and specificity (Sp) at 85%. The required numbers of disease-positive (N_pos) and disease-negative (N_neg) samples were approximated using the standard normal approximation for estimating a proportion with a desired precision:

$$N = \frac{Z^2_{1-\frac{\alpha}{2}} \times p(1-p)}{d^2}$$

where $Z_{1-\alpha/2}$=1.96 for a two-sided 95% confidence interval, p represents the expected Se or Sp, and d is the allowable absolute error (half-width of the CI). Using d=7.5%, the estimated minimum sample size is approximately 62 pancreatic cancer cases (for Se = 0.90) and 87 controls (for Sp = 0.85), resulting in a total cohort size of ~149 samples to achieve ±7.5% precision for both sensitivity and specificity.

We assembled an expanded discovery cohort comprising 63 pancreatic cancer cases, 50 healthy controls, and 50 disease controls (pancreatitis), which meets the calculated requirement under the ±7.5% error criterion. The revised manuscript and Supplementary Information now clearly describe the enlarged cohort and EnCODE performance on the discovery cohort. To further assess generalizability, we tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15). EnCODE showed consistent performance between the discovery and external validation cohorts, supporting its reproducibility and generalizability across medical centers (**Figure 5, Supplementary Figure 16**).

In parallel, we performed head-to-head comparisons with qRT-PCR. Mechanistically, qRT-PCR/dPCR rely on fluorescence-based amplification and/or compartmentalization, whereas

EnCODE couples biochemical color generation with a mathematically defined, reversible encoding–decoding framework, leveraging experimentally validated spectral linearity and RGB additivity to map multi-marker information into a structured colorimetric space. Practically, PCR-based multiplex panels often require multiple reactions, wells, or cartridges as the number of targets increases, followed by post hoc aggregation of Ct values or droplet counts. By contrast, EnCODE performs single-tube multiplex decoding in one reaction, integrates multiple biomarker contributions through controlled enzymatic color mixing, and can directly output a composite risk score without requiring a multi-reaction workflow. This design reduces assay complexity scaling with panel size and mitigates tube-to-tube or well-to-well variability introduced by separate amplification reactions.

We additionally performed parallel benchmarking experiments. Starting from pre-amplification step, we ran the full workflow and compared EnCODE-derived abundance estimates against qRT-PCR quantification for matched targets. EnCODE showed strong concordance with qRT-PCR measurements ($R^2 > 0.99$), supporting that EnCODE can achieve comparable analytical performance while providing complementary advantages in multiplex integration, information-dense encoding, and direct score readout (**Supplementary Figure 13 a-c**). Taken together, these revisions clarify both the novelty of the encoding paradigm and its potential translational utility as a multiplex, scalable colorimetric framework. (**Line 454-466, Figure 5, Supplementary Figure 16, Line 406-413 and Supplementary Figure 13**)

Revised Main Text (Line 454-466):

Finally, we evaluated EnCODE using a clinical serum cohort comprising pancreatic cancer (n = 63), benign pancreatic disease controls (pancreatitis, n = 50), and healthy controls (n = 50) (**Figure 5f, Supplementary Table 4**). EnCODE generated directly interpretable prediction scores and enabled robust classification of pancreatic cancer versus controls, achieving an AUC (area under the ROC curve) of 0.9586 (95% CI: 0.9310–0.9862), with 96.83% sensitivity (95% CI: 89.14–99.44%) and 86.00% specificity (95% CI: 77.86–91.47%). Importantly, EnCODE provide effective discrimination from benign pancreatic disease (pancreatitis) (**Figure 5g–j**). In parallel, qRT-PCR profiling of the same miRNA panel recapitulated the expression differences in clinical samples and showed comparable discrimination performance, supporting the analytical consistency between EnCODE and conventional PCR-based workflows (**Supplementary Figure 16b–d**). To further assess generalizability, we tested an independent external validation cohort collected at two additional medical centers (n = 50; pancreatic cancer n = 20, pancreatitis n = 15, healthy n = 15), which yielded consistent classification accuracy (**Supplementary Table 5, Figure 5k, l**).

Revised Figure 5:

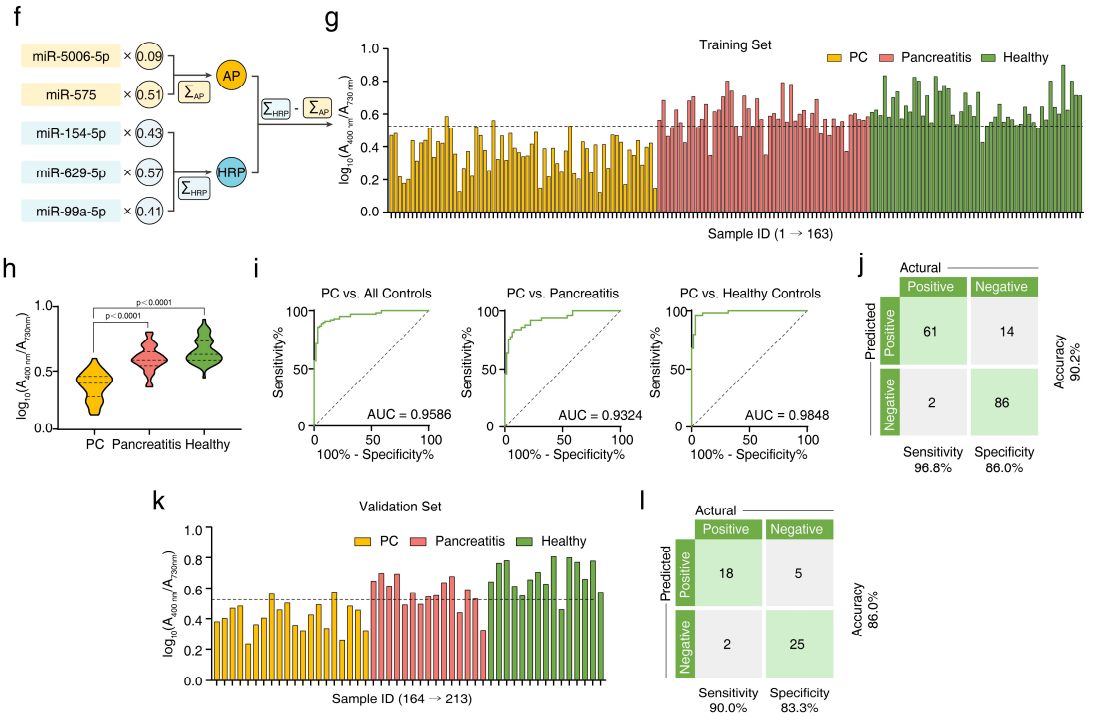
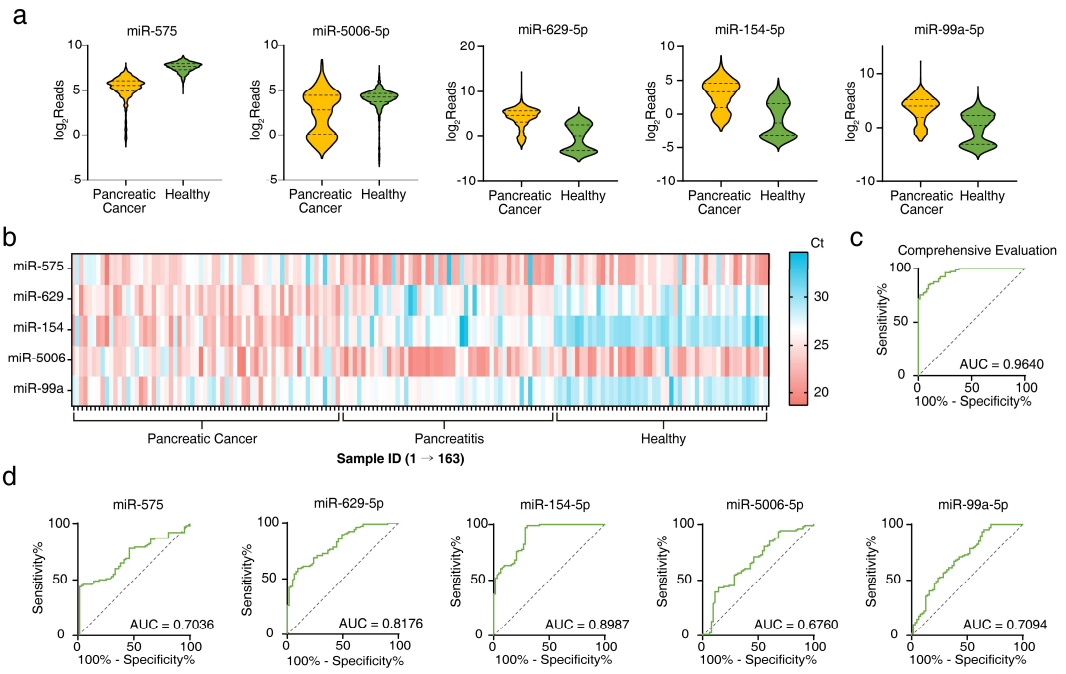


Figure 5. (f) Schematic diagram illustrates the pancreatic cancer prediction model: Score = 0.43×miR-154-5p + 0.57×miR-629-5p + 0.41×miR-99a-5p – (0.09×miR-5006-5p + 0.51×miR-575). (g) Representative EnCODE readouts for 163 clinical serum samples in the expanded cohort, including 63 pancreatic cancer cases, 50 pancreatitis controls, and 50 healthy controls. (h) Risk-score–based prediction results for pancreatic cancer classification derived from EnCODE measurements in the expanded clinical cohort. (i) ROC curve of EnCODE performance for clinical sample classification. (j) Confusion matrix summarizing classification accuracy of EnCODE in the expanded clinical cohort. (k) Representative EnCODE readouts for an independent external validation cohort of 50 clinical serum samples collected from additional medical centers, including 20 pancreatic cancer cases, 15 pancreatitis controls, and 15 healthy controls. (l) Confusion matrix summarizing EnCODE classification accuracy in the external validation cohort. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Statistical analysis was performed using a Student's t-test.

Revised Supplementary Figure 16:

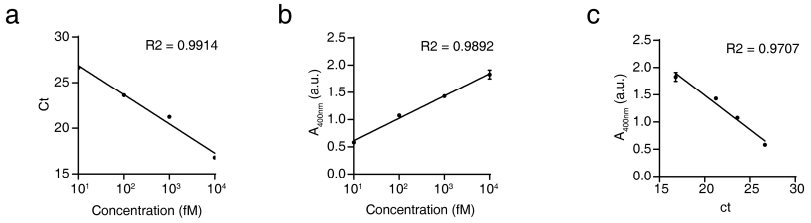


Supplementary Figure 16. Diagnostic model and PCR-based performance benchmarking. (a) Differential expression analysis of the five selected miRNA markers in the GSE211692 serum dataset. (b) miRNA expression profiles measured in the expanded clinical cohort, including 63 pancreatic cancer serum samples, 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.

Revised Main Text (Line 406-413):

To further assess metrological stability in a workflow-relevant setting, we benchmarked EnCODE using miRNA samples processed through the same pre-amplification step as in the clinical workflow. Across the tested concentration series, EnCODE readouts closely followed the input miRNA levels and showed strong agreement with qRT-PCR quantification, supporting reliable quantitative transfer from miRNA abundance to enzymatic output (**Supplementary Figure 13a–c**). In parallel, both the HRP–ABTS and AP–pNPP systems exhibited robust linear relationships between miRNA concentration and colorimetric features at the levels of absorbance and RGB values, further supporting a stable miRNA-to-enzyme signal mapping (**Supplementary Figure 13d-e**).

Revised Supplementary Figure 13:



Supplementary Figure 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.

The work does not adequately support its major conclusions and claims. The 32-sample clinical validation is insufficient for reliable diagnostic performance estimates, requiring independent validation cohorts and multi-site studies. Claims of "Polaroid-like simplicity" contradict the multi-step, 7-hour protocol described, and no evidence supports visual interpretation by non-specialists. Assertions of superior performance over established methods lack supporting comparative data, and the biomarker selection process from existing datasets may involve circularity between discovery and validation.

Response:

Thank you for pointing out potential inappropriate wording and conclusions in the manuscript. In the revised manuscript, we have removed or reworded statements that may have implied clinical readiness or oversimplified the workflow. Specifically, references to "Polaroid-like simplicity" and visual interpretation by non-specialists have been eliminated. We now provide an accurate description of the multi-step, 7-hour protocol, while clearly indicating which components of the workflow could potentially be automated in the future. Furthermore, we emphasize that the novel contribution of this work lies in the biochemical-mathematical encoding framework, which enables structured multiplexed color detection.

To minimize operator-dependent variation during image-based color readout, all sample handling and EnCODE procedures were performed under environmental chamber (25 ± 1 °C and 50 ± 5% relative humidity). For image-based readout, we additionally standardized the imaging geometry using a fixed-distance holder and employed uniform diffuse illumination (LED ring light or a backlight lightbox) to minimize shadows and spatial non-uniformity. Each assay run included internal reference standards for plate-to-plate normalization, thereby correcting potential variations in spectral response and enzyme activity. In practical deployment, we recommend reporting results primarily via objective RGB/spectral decoding (with visual scoring as an auxiliary output) to further reduce subjectivity.

To quantitatively evaluate the variability of human visual interpretation, three blinded raters (n = 3) independently scored the same set of representative outputs spanning four intensity levels (strong positive, weak positive, strong negative, and weak negative). The interpretation results were highly consistent across the three raters, indicating that subjective judgment is unlikely to introduce major variability in result interpretation. These results, together with the standardized imaging and objective RGB/spectral decoding workflow, support that EnCODE readouts can be reliably interpreted with improved consistency (We assembled an expanded discovery cohort comprising 63 pancreatic cancer cases, 50 healthy controls, and 50 disease controls (pancreatitis), which meets the calculated requirement under the ±7.5% error criterion. The revised manuscript and Supplementary Information now clearly describe the enlarged cohort and EnCODE performance on the discovery cohort (**Supplementary Table 1**).

We also appreciate your comments on the biomarker discovery and validation process. Specifically, GSE211692 (serum cohort, n=16190) and GSE163031 (tissue cohort, n =15) were used for biomarker screening and model construction, where we retained miRNAs showing consistent

dysregulation directionality across serum and tissue to enhance biological plausibility. Importantly, the final miRNA panel and prediction model were further evaluated on an independent external validation cohort, GSE106817(n=4046), which was processed and analyzed separately from the training procedure to avoid any information leakage between training and validation steps. The consistent performance observed in this independent dataset supports the robustness and generalizability of the proposed miRNA marker panel.

These updates have been incorporated into the revised manuscript and supplementary materials (**Line 35-37, Line 77-81, Line 289-291, Line 557-558, Line 302-306, Line 653-657, Supplementary Table 1 and Line 431-444**).

Revised Main Text (Line 35-37):

We enhanced data density through three-dimensional color encoding and implemented hyperspectral imaging-based data analysis, thereby facilitating an intuitive color-coded molecular readout.

Revised Main Text (Line 77-81):

This innovation couples enzymatic colorimetric encoding with quantitative decoding, enabling multiplex biomarker integration with instrument-light readout and scalable workflows. These features support a feasible route toward decentralized deployment, particularly when paired with standardized imaging, internal calibration, and automated analysis.

Revised Main Text (Line 289-291):

Direct classification of samples is achieved through standardized spectral decoding or quantitative color-feature analysis (**Figure 3i**).

Revised Main Text (Line 557-558):

Spectral detection enables quantitative decoding of the information encoded in colors, while visual color readout provides an intuitive and instrument-light option for rapid assessment.

Revised Main Text (Line 302-306):

To quantify the variability of subjective visual interpretation, three blinded raters independently classified representative EnCODE outputs spanning strong/weak positive and strong/weak negative cases, showing substantial inter-rater agreement (accuracy = 94.4%; **Supplementary Table 1**). This further supports the consistency of visual readout when complemented by standardized imaging and objective decoding.

Revised Main Text (Line 653-657):

All assays were conducted under controlled ambient conditions (25 ± 1 °C, 50 ± 5% relative humidity). For image acquisition, a fixed-distance holder was used to standardize the camera-to-plate geometry, together with uniform diffuse illumination (LED ring light or backlight lightbox) to reduce shadows and spatial inhomogeneity. Each run included standardized reference wells for plate-to-plate calibration.

Revised Supplementary Table 1:

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

Sample	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

Revised Main Text (Line 431-444):

After determining the clinical testing workflow, we applied EnCODE to pancreatic cancer diagnosis and prediction. We first attempted to construct a pancreatic cancer diagnosis prediction model. Using the GSE211692 dataset (included serum miRNA expression levels from healthy controls and pancreatic cancer patients) and the GSE163031 dataset (included pancreatic tissue miRNA expression levels from normal and cancerous samples) from the Gene Expression Omnibus (GEO) database, we identified three upregulated and two downregulated miRNAs consistently expressed across both sample types (**Figure 5c, Supplementary Figure 16a**). Subsequently, a pancreatic cancer diagnosis model was constructed via binary logistic regression: Score = 0.43 × miR-154-5p + 0.57 × miR-629-5p + 0.41 × miR-99a-5p − (0.09 × miR-5006-5p + 0.51 × miR-575). Notably, initial miRNA concentrations were log10-transformed to make the concentration data compatible with the LATE-PCR amplification mechanism. The model demonstrated strong discrimination between pancreatic cancer and healthy samples in public datasets (**Figure 5d**) and performed robustly across serum (GSE106817, GSE211692) and tissue (GSE163031) cohorts (**Figure 5e**).

Several significant flaws in data analysis, interpretation, and conclusions prohibit publication in the current form. Statistical issues include inadequate sample size for diagnostic performance claims, wide confidence intervals indicating high uncertainty, potential overfitting with five variables and only twelve positive cases, and unclear validation methodology. Analytical concerns encompass missing validation of linearity under real-world conditions, insufficient assessment of matrix effects, limited stability data, and no inter-operator validation. The interpretation problems include claims exceeding supporting evidence, premature clinical readiness assertions, and oversimplified presentation of the complex protocol.

The methodology falls below expected standards for diagnostic test evaluation. The clinical study lacks proper power analysis and adequate controls, biomarker selection may involve methodological circularity, and critical validation studies for stability, reproducibility, and specificity are missing.

Environmental variables affecting colorimetric detection are insufficiently controlled, and potential confounding factors receive limited assessment. While the basic chemistry appears sound, the clinical validation methodology is inadequate for the diagnostic claims presented.

Response:

We sincerely thank you for the rigorous and detailed critique. In the revised manuscript, we have comprehensively addressed these concerns by expanding the clinical evaluation, clarifying the validation methodology, extending robustness testing under more realistic conditions, and explicitly repositioning the study as a methodological proof-of-concept to avoid over-interpretation of clinical readiness.

First, we agree that the original 32-sample cohort was underpowered for reliable estimation of diagnostic performance. We assembled an expanded discovery cohort comprising 63 pancreatic cancer cases, 50 healthy controls, and 50 disease controls (pancreatitis), which meets the calculated requirement under the ±7.5% error criterion. The revised manuscript and Supplementary Information now clearly describe the enlarged cohort and EnCODE performance on the discovery cohort (**Figure 5, Supplementary Figure 16**).

Second, to mitigate concerns regarding uncertainty and variability, we now report error bars throughout the revised Results and Supplementary Information. Under standardized acquisition conditions, spectral and RGB readouts showed low well-to-well and batch-to-batch variability, with coefficients of variation typically below 10% for key readouts. For samples sharing the same composite score with matched component concentrations, chromatic features (RGB/Hue) remained tightly distributed; in contrast, when different underlying component combinations produced the same score, Hue distributions broadened as expected from the geometry of color-space mapping, but without impairing score-based classification.

Third, we measured the target miRNA samples using EnCODE. Consistent with the clinical-sample workflow, the target miRNAs underwent the same pre-amplification step prior to EnCODE readout. Across the tested range (10 fM to 10 pM), both the spectral readout and the colorimetric (RGB-based) readout showed strong linear relationships with the logarithm of miRNA abundance, with $R^2 > 0.98$ for the spectral features and $R^2 > 0.97$ for the color features. These results support a robust linear mapping between EnCODE optical signals and miRNA concentration over a 10 fM–10 pM dynamic range (**Supplementary Figure 13**).

Further, we strengthened matrix-effect and robustness assessments by including interference testing (e.g., hemolysis, lipemia, high-protein samples, anticoagulants/coagulants, and common pancreatic cancer therapeutics) as well as storage-stability evaluations across typical temperatures and handling windows, providing actionable pre-analytical guidance (**Supplementary Figure 14, Supplementary Figure 15**). Inter-operator robustness was further evaluated by three independent operators who conducted the full EnCODE workflow on the same sample set, and the results showed low operator-dependent variability, supporting assay reproducibility (**Supplementary Figure 10**).

We also appreciate your comments on the biomarker discovery and validation process. Specifically, GSE211692 (serum cohort, n=16190) and GSE163031 (tissue cohort, n =15) were used for biomarker screening and model construction, where we retained miRNAs showing consistent dysregulation directionality across serum and tissue to enhance biological plausibility. Importantly, the final miRNA panel and prediction model were further evaluated on an independent external validation cohort, GSE106817(n=4046), which was processed and analyzed separately from the

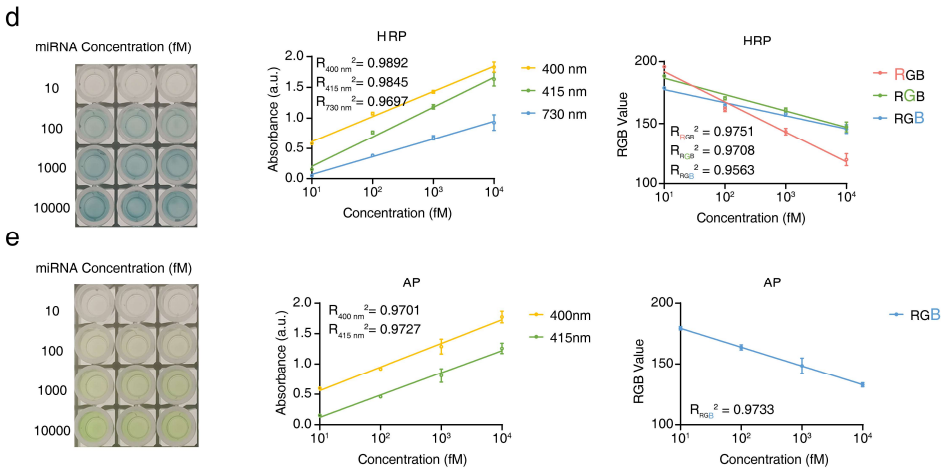
training procedure to avoid any information leakage between training and validation steps. The consistent performance observed in this independent dataset supports the robustness and generalizability of the proposed miRNA marker panel.

Finally, we agree that parts of the original narrative overstated simplicity and clinical readiness. Accordingly, we removed "Polaroid-like" language and any implication of non-specialist visual diagnosis, and we revised the manuscript to accurately reflect the multi-step workflow while explicitly indicating which steps are amenable to future automation. We now emphasize that the main contribution is the biochemical–mathematical encoding/decoding framework and its analytical validation, with clinical results presented as a proof-of-concept demonstration of feasibility rather than a finalized diagnostic claim. **(Line 406-413, Supplementary Figure 13, Line 425-430, Supplementary Figure 15, Line 414-423, Supplementary Figure 14, Line 306-310 and Supplementary Figure 10)**

Revised Main Text (Line 406-413):

To further assess metrological stability in a workflow-relevant setting, we benchmarked EnCODE using miRNA samples processed through the same pre-amplification step as in the clinical workflow. Across the tested concentration series, EnCODE readouts closely followed the input miRNA levels and showed strong agreement with qRT-PCR quantification, supporting reliable quantitative transfer from miRNA abundance to enzymatic output (**Supplementary Figure 13a–c**). In parallel, both the HRP–ABTS and AP–pNPP systems exhibited robust linear relationships between miRNA concentration and colorimetric features at the levels of absorbance and RGB values, further supporting a stable miRNA-to-enzyme signal mapping (**Supplementary Figure 13d-e**).

Revised Supplementary Figure 13:



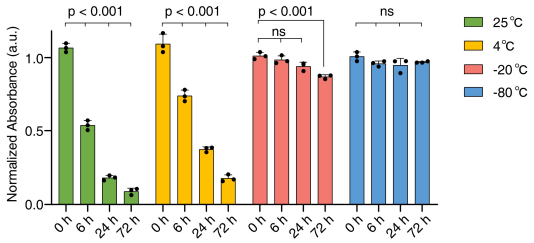
Supplementary Figure 13. (d) Performance validation of EnCODE for miRNA detection using the HRP–ABTS system. Left: representative photographs of EnCODE colorimetric outputs at different miRNA concentrations; middle: relationship between absorbance at characteristic spectral peaks and miRNA concentration; 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).

Revised Main Text (Line 425-430):

In addition, we systematically assessed the effects of storage conditions (4°C, -20°C, -80°C) and storage duration (0–72 hours) on the accuracy of miRNA detection. Our results showed that miRNA signals remained stable when samples were stored at -20°C for up to 1 day or at -80°C for up to 3 days, with no significant deviation from fresh controls. These findings provide clear guidance on sample storage conditions for reliable EnCODE analysis (**Supplementary Figure 15**).

Revised Supplementary Figure 15:

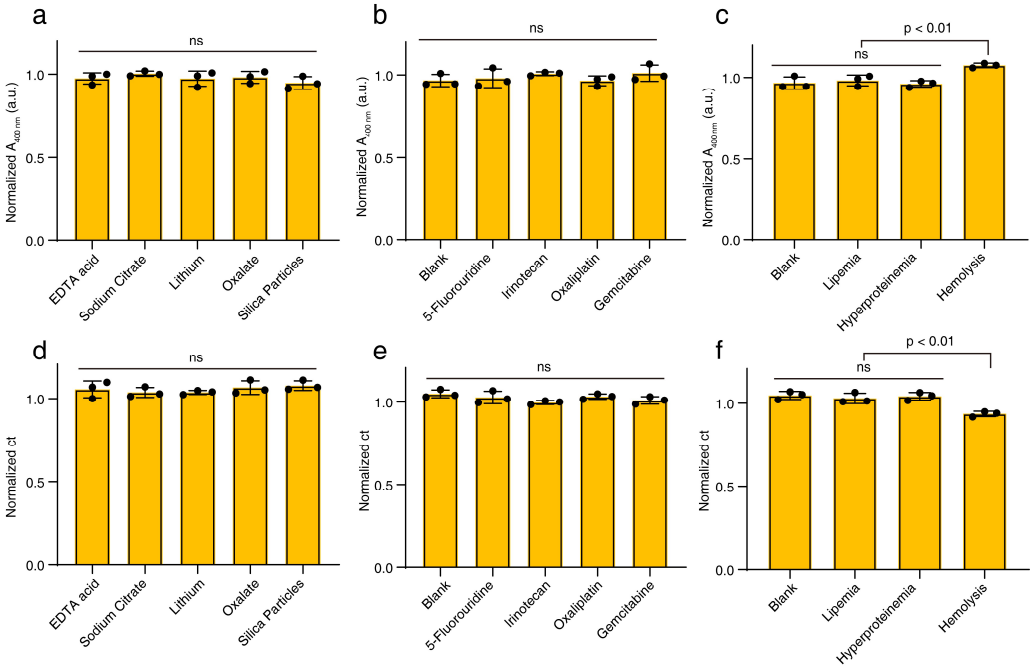


Supplementary Figure 15. Impact of sample storage conditions and duration on miRNA detection results using EnCODE. Data are presented as the mean ± s.d. from three independent experiments (n = 3). Statistical analysis was performed using a Student's t-test, and p values are indicated.

Revised Main Text (Line 414-423):

To evaluate robustness under clinically relevant conditions, we assessed common pre-analytical interferences and therapeutic agents using aliquots from the same-source blood sample. We tested anticoagulants/coagulants, representative pancreatic cancer drugs (5-Fluorouridine, Irinotecan, Oxaliplatin, and Gemcitabine), and simulated matrix interferences (lipemia, high protein, and hemolysis) by adding equal volumes of Intralipid, albumin, or red-blood-cell lysate, respectively. Anticoagulants/coagulants, drugs, and lipemic/high-protein conditions showed no appreciable impact on EnCODE readouts, whereas hemolysis caused a modest upward shift consistent with erythrocyte miRNA release (**Supplementary Figure 14a-c**). Notably, EnCODE readouts followed the same directional trends as qRT-PCR on the matched samples, supporting that these sample conditions do not materially bias miRNA quantification (**Supplementary Figure 14d-f**).

Revised Supplementary Figure 14:

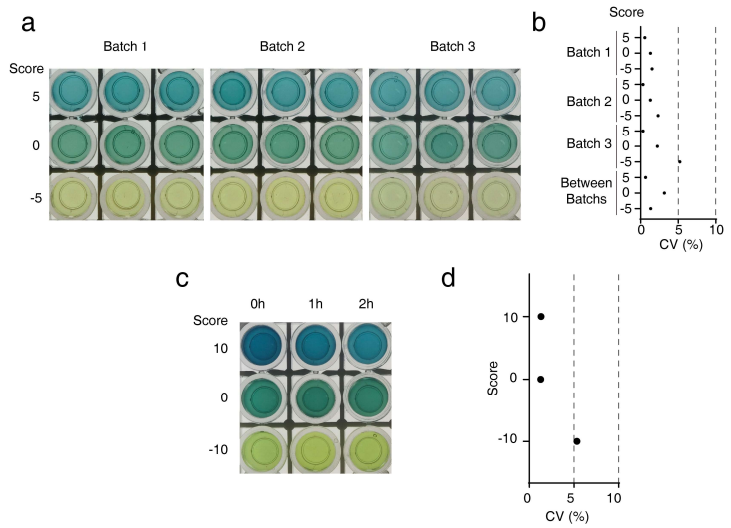


Supplementary Figure 14. Interference validation of the EnCODE platform. (a) EnCODE miRNA readouts for blood aliquots collected under different anticoagulant/coagulant conditions, reported as normalized absorbance. (b) EnCODE miRNA readouts (normalized absorbance) for matched blood aliquots with or without 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 Student's t-test, and p values are indicated.

Revised Main Text (Line 306-310):

To further assess robustness, we evaluated inter-batch reproducibility and examined the impact of pre-extraction room-temperature holding time. Both intra-plate and inter-batch variations remained low, and the decoded scores showed no appreciable drift within the tested handling-time window (**Supplementary Figure 10**).

Revised Supplementary Figure 10:



Supplementary Figure 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 (d) across different holding times and risk-score levels.

Critical details necessary for reproduction are missing throughout the methods section. These include enzyme standardization and quality control procedures, environmental control requirements, specific protocols for visual interpretation, statistical analysis software parameters, detailed preparation protocols for DNA templates and conjugates, and storage requirements for reagents. Importantly, despite stating that code is available "upon reasonable request," no data analysis scripts, algorithms for spectral decoding, or computational workflows are provided or deposited in public repositories. This lack of computational transparency significantly hampers reproducibility, particularly given the mathematical encoding/decoding framework central to the method. Although supplementary methods provide some additional detail, key aspects remain insufficiently described for independent reproduction.

Response:

We thank you for emphasizing the importance of methodological transparency and reproducibility.

In response, we have substantially expanded the Methods section and Supplementary Information to address all points raised. In addition, we now provide the complete raw datasets underlying all reported experiments. To ensure computational transparency, we have also provided the data-processing scripts in the Supplementary Information as part of the revised submission package. Together, these additions substantially improve reproducibility and enable independent researchers to replicate both the experimental procedures and the analytical workflow underlying the EnCODE framework. (**Line 633-635, Line 653-657, Line 772-789, Line 659-670, Line 623-624, Line 677-689, Line 173-179 and Supplementary Code**)

Revised Main Text (Line 633-635):

HRP and AP were stored at 4 °C in PBS containing 50% (v/v) glycerol. The reaction buffer consisted of 500 mM Tris–HCl (pH 8.0), 100 mM $MgCl_2$, 0.1 mM $ZnCl_2$, and 100 mM KCl. Diluted enzyme solutions were stable at 4 °C for up to 1 week.

Revised Main Text (Line 653-657):

All assays were conducted under controlled ambient conditions (25 ± 1 °C, 50 ± 5% relative humidity). For image acquisition, a fixed-distance holder was used to standardize the camera-to-plate geometry, together with uniform diffuse illumination (LED ring light or backlight lightbox) to reduce shadows and spatial inhomogeneity. Each run included standardized reference wells for plate-to-plate calibration.

Revised Main Text (Line 772-789):

The LIMMA package from the Bioconductor project in R software[52] was utilized to identify differentially expressed miRNAs. Given our objective to analyze serum miRNAs for subsequent detection, we aimed to find miRNAs with greater expression differences in serum samples. The cutoff for differentially expressed miRNAs in serum samples from the GSE211692 dataset was set as log2 fold change (FC) ≥ 3.5 or log2 fold change (FC) ≤ -2.5 and an error discovery rate (FDR) ≤ 0.05, while in the tissue samples from the GSE163031 dataset, the cutoff was set as |log2 fold change (FC)| ≥ 1 and an FDR < 0.05. Overlapping miRNA expression profiles between the GSE211692 and GSE163031 datasets were retained for further study.

Binary logistic regression was used to construct a probability model for pancreatic cancer using the selected miRNAs as predictors. Prior to model fitting, miRNA concentrations were log10-transformed to reduce skewness and improve numerical stability across the wide dynamic range introduced by the pre-amplification step. Regression coefficients were estimated by maximum likelihood using the standard iterative algorithm implemented in SPSS (Version 20). A parsimonious model was derived using forward stepwise selection in SPSS (Forward: Wald), with candidate variables sequentially entered according to prespecified significance criteria (*entry* $p < 0.05$). Model discrimination was evaluated by receiver operating characteristic (ROC) analysis and summarized by the area under the curve (AUC). The final score was reported as a linear combination of the retained miRNAs with their estimated coefficients and directions of effect.

Revised Main Text (Line 659-670):

The template DNAs used in this study were custom-synthesized and purified by Huzhou Hippo Biotech (Huzhou, China) and Sangon Biotech (Shanghai, China). Purification was performed by HPLC, and mass spectrometry confirmed a single dominant peak.

Initially, combine 10 μM of linear template DNA with 20 μM of the target. Subsequently, subject this mixture to heat annealing by incubating it at 95 °C for 3 minutes, followed by a gradual cooling to 12 °C at a rate of 1 °C per minute to form a linear DNA template-target heterodimer. Next, incubate the 2 μM linear template DNA-primer heterodimer with 40 U/μL SplintR Ligase in 1× SplintR Ligase buffer at 16 °C for a minimum of 1 hour to yield a circular template DNA-primer heterodimer. To confirm the successful circularization of the template DNA, treat the system with 0.2 U/μL of Deoxyribonuclease I and 2 U/μL of Deoxyribonuclease III in 1× appropriate buffer at 37 °C for 1 hour to completely remove the primers and linear template. Verify the results through

polyacrylamide gel electrophoresis (PAGE).

Revised Main Text (Line 623-624):

All reagents are stored strictly according to the supplier's instructions. Nucleic acids are stored at -20°C.

Revised Main Text (Line 677-689):

Initially, a 30 μM solution of enzyme (Amylase or Horseradish Peroxidase) was combined with 1.5 mM of the bifunctional crosslinker NHS-PEG4-N3 in 1× PBS buffer. The mixture was gently stirred for 1 hour at 300 rpm and 25 °C to convert the amino groups of the enzyme into azide groups. Once the reaction was complete, the enzyme was purified using a 30 kDa ultrafiltration centrifuge tube and washed 6 times with 1× PBS buffer (pH 7.4) to remove excess NHS-PEG4-N3. Subsequently, the azide-modified enzyme was mixed with DBCO-modified DNA barcodes at a concentration ratio of 10 μM:100 μM and incubated at 4 °C for 24 hours to produce the enzyme-DNA barcode conjugates. To eliminate excess DBCO-DNA, the product was filtered again using a 30 kDa ultrafiltration centrifuge tube and washed with 1× PBS buffer (pH 7.2). Finally, the conjugated enzymes were then stored in PBS containing 50% glycerol at -20 °C for future use.

All conjugates were verified for successful enzyme-DNA conjugation using SDS-PAGE to confirm the expected molecular weight shift. Enzyme activity was assessed using HRP or AP activity assays to ensure that the conjugates retained sufficient enzymatic activity.

Revised Main Text (Line 173-179):

The absorption spectrum of a single enzyme solution scales linearly with enzyme concentration, enabling theoretical prediction of spectra of mixed solutions based on reference spectra of blank substrate, HRP (concentration A), and AP (concentration B). The calculation formula is as follows:

$$Absorption_{mixed} = \frac{X}{A}\left(Absorption_{HRP} - Blank\right) + \frac{Y}{B}\left(Absorption_{AP} - Blank\right) + Blank$$

where X and Y represent the concentration of HRP and AP in the mixed solution (**Figure 2m**). By simplifying spectra to the three critical wavelengths of 400 nm, 415 nm, and 730 nm, we efficiently decoded the concentrations of HRP and AP in the mixed solution (**Figure 2n**). Additionally, the RGB values can also be used for decoding, with the same resolution capability (**Figure 2o**).

The manuscript requires major revision before publication consideration. The core technical approach has merit as a proof-of-concept study, but clinical claims are premature and inadequately supported. The authors should reframe this as a methods development study rather than a clinical diagnostic platform, conduct properly powered clinical validation with independent cohorts, provide comprehensive reproducibility and stability data, include direct comparisons with established detection methods, make all analysis code and algorithms publicly available, and correct the numerous presentation errors throughout the manuscript. Additionally, the pervasive spelling and grammatical errors, including technical terms in figures, require professional editing to meet publication standards.

With appropriate scope and validation, this work could potentially contribute to the biosensing literature. However, substantial additional evidence and methodological improvements are needed

to support the current conclusions and clinical claims.

Response:

We sincerely thank you for their thoughtful and constructive feedback. We agree that further evidence and methodological refinements are necessary to strengthen our conclusions. In the revised manuscript, we have expanded the scope of our study by increasing the sample size, adding independent validation cohorts, and providing comprehensive data on reproducibility, stability, and specificity. We have also clarified the methodological framework, focusing on the biochemical-mathematical encoding-decoding system rather than premature clinical claims. These updates aim to provide a clearer and more robust foundation for the EnCODE platform's potential contribution to the biosensing literature. We believe these revisions address your concerns and significantly improve the manuscript's clarity and rigor.

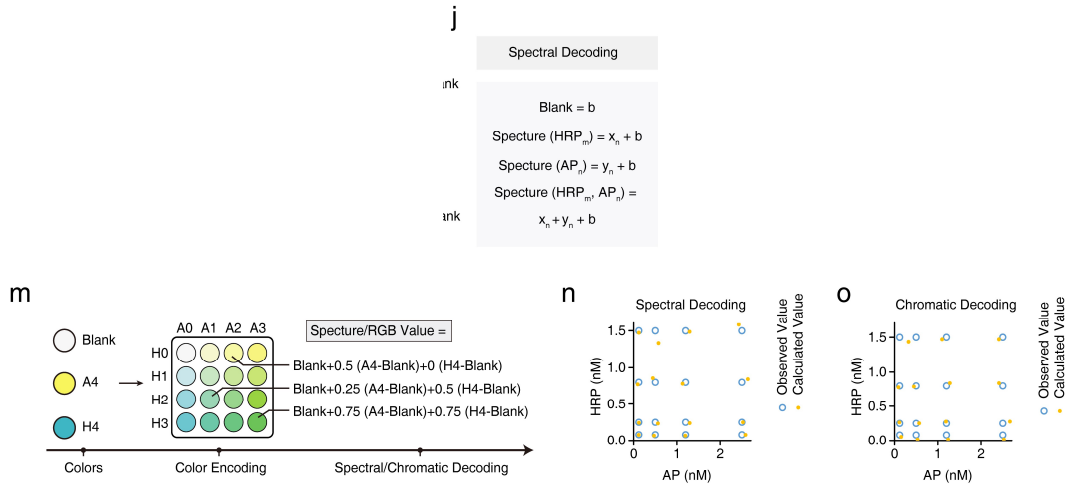
Minor Issues

The manuscript also needs professional editing - there are numerous spelling errors and grammatical mistakes throughout, including in figure labels (e.g., "spectural" instead of "spectral", "competitive binging" instead of "binding").

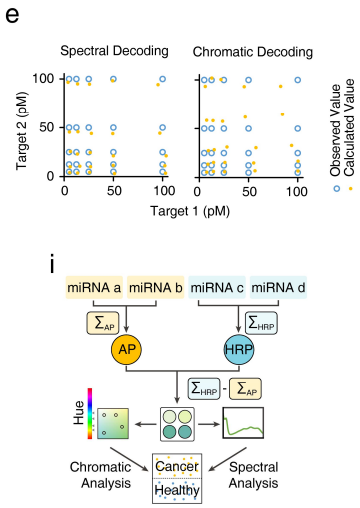
Response:

We sincerely appreciate your comment regarding the need for professional editing of the manuscript. We have thoroughly reviewed the manuscript and made corrections to all identified spelling and grammatical errors. Additionally, we ensured consistency in technical terms and refined the overall language to meet the required standards for publication. (**Figure 2, Figure 3, Figure 4, Figure 6**)

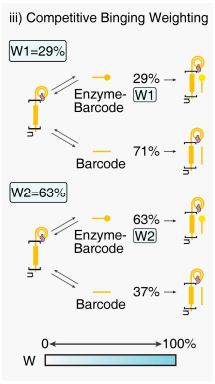
Revised Figure 2:



Revised Figure 3:



Revised Figure 4:



Revised Figure 6: