

Detectors convert transient RNA sequences into stable DNA barcodes for high-throughput analysis of RNA-dependent processes

Corresponding Author: Dr Seth Shipman

Parts of this Peer Review File have been redacted as indicated as we could not obtain permission to publish the reports from one of the peer reviewers.

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)

This manuscript describes a novel and broadly useful “Detector” that enables conversion of RNA signals into a more durable DNA output signal in bacteria. Detectors operate through a toehold switch-inspired regulatory mechanism where a 5' hairpin structure, which initially blocks retron reverse transcription, is unwound by a complementary trigger RNA, promoting RNA refolding and activation of retron reverse transcription. The authors demonstrate that Detectors can be used to detect natural RNAs (e.g. Zika virus fragments, phage transcripts) to induce production of barcoded DNA reporters, providing longer-lived products that can be recorded through downstream sequencing. The manuscript provides a detailed study of Detector design parameters based on a library of ~7800 analyzed variants and further uses machine learning to identify the most salient system features. They confirm that the system can work in multiple bacterial hosts and demonstrate that phage-specific Detectors can be used to record the preferences of phage infection across a population of different strains of bacteria.

Overall, the manuscript is very clearly written, the results support the robust operation of the system, and the testing and optimization based on a high-throughput library of Detectors goes well beyond the norm for the first paper describing a new RNA-based regulatory system. The Detectors themselves represent a valuable new capability for synthetic biology, which could be deployed for a multitude of different applications where long-term storage of transient RNA information is needed. For publication in a high-impact journal like Nature Biotechnology, the main weakness of the paper is the deployment of Detectors for phage-host interactions, where experiments are limited to three hosts or phages and it is unclear how much the method improves over more conventional approaches. I recommend publication once the points below are addressed:

1. The authors should provide more details about the plasmid design for the Detector system. Critical components of the Detector plasmid are not sufficiently described, such as the origin of plasmid, the RT expression, etc. Those details are essential for readers to understand and reproduce the experiments.
2. It is unclear in the manuscript how the Detector approach improves over other approaches that could be used to examine phage-host interactions. For instance, could DNA sequencing of surviving cells or phage populations after infection be used to examine phage-host interactions with similar outcomes? The need to supply chloramphenicol to suppress cell lysis seems as if it could limit the approach when applied to more diverse bacterial communities for phage-host studies. Providing a clearer and more detailed description of the benefits and limitations of the Detectors would be valuable for readers.
3. The first sentence of the Results section is confusing and should be revised. It currently suggests that the retron ncRNA also contains a single-stranded DNA region that is also reverse transcribed; the timing of the conversion of msd RNA to DNA should be clarified.
4. It would be helpful to have arrowheads indicating the 3' ends of the RNAs/DNAs in the figures. This is particularly true in Fig. 1a where there is the trigger, retron RNA, and nascent ssDNA (which confusingly is also the same color as the a1/a2

region).

5. Did the authors test any Detectrons with loop lengths shorter than 12 nt or are there system features (e.g. a1/a2 lengths) that prevented this? The ML results and RNA folding considerations suggest that shorter loops could lead to more stable structures and lower signal leakage. More information should be provided as to why certain ranges of domain lengths were selected.
6. Some additional information concerning the Detectron library would be helpful to state in the main text. Was the a1/a2 length varied? Did all elements in the library use the same trigger sequence? Were the stem2 sequences randomized?
7. A Detectron where the trigger invades into stem2 would probably perform better since it would discourage reversal of the strand-displacement reaction and thus more stable a1/a2 binding upon activation. Were any systems in the library designed so that the trigger invaded into stem2? If not, it may be interesting to revisit the library testing results to see if partial stem2 strand invasion (e.g. 1-3+ base overlap with the trigger) led to any performance improvements.
8. Toehold-mediated RNA refolding and stem-formation mechanisms similar to those used in the Detectrons have been used in a few papers (e.g. PMID: 39127751, 39036067, 41272150). These earlier papers should be cited.
9. For Fig. 1h, the signal from cells with Detectron alone is higher than Detectron + scrambled RNA, which is unexpected and perhaps a statistically significant difference. Is the increased signal due to the scrambled RNA interacting with the Detectron to inhibit its activation?
10. The ON/OFF ratio for T7-V2 exceeds 30 in Figure 4 but is reduced to 2 in Figure 5c. Please explain why this large drop occurs.
11. For the machine learning model, the authors should provide more details in how they define the SHAP baseline for both feature importance score and ON/OFF ratio across different feature distributions.
12. In Figure 3b, only the test set data are shown, but the caption mentions the training and test set data being present. Please revise.
13. Page 16: "onT7" should be "on T7".
14. In Extended Data Fig. 1, the priming G site should be located closer to the 5' end rather than the 3' end in the ECO1 ncRNA.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

This manuscript by Han and Shipman introduces "Detectrons," a synthetic biology platform that couples programmable RNA sensing via toehold switches with retron-mediated DNA barcode generation. By converting transient RNA signals into stable DNA outputs, the authors propose a new framework for multiplexed sensing and recording of transcriptional events. The work is technically sophisticated, combining circuit design, high-throughput screening, and machine learning, and is applied to an interesting use case, namely phage-host interaction profiling in pooled bacterial cultures.

The authors position Detectrons as the first RNA-to-DNA transduction system. While the implementation is elegant, the conceptual distinction from prior molecular recording systems (e.g., retron-based recording or CRISPR-based acquisition systems) is not fully articulated, particularly in the Introduction. The manuscript would benefit from a clearer and more explicit comparison with recent retron-based recording approaches and RNA-guided recording strategies. As it stands, the advance appears at least in part incremental—i.e., based on the integration of existing components (toehold switches and retrons)—rather than entirely conceptual. It would therefore be important for the authors to better define what fundamentally new capability is enabled by Detectrons and in which aspects this system clearly outperforms existing approaches.

Overall, the manuscript is well organized and the experimental workflow is clearly described. However, several aspects require clarification and, in some cases, additional support to fully justify the scope of the claims and strengthen the overall impact.

Major comments

1. The hypothesis that sequestration of the a2 region inhibits reverse transcription and that strand displacement restores activity is supported by the experimental data, although no direct structural or biochemical validation is provided. Given that the observed improvements in ON/OFF ratios are attributed to structural parameters (stem1, stem2, loop, and bulge), it is unclear whether the effect of toehold length itself was systematically investigated. In particular, is the length of the region highlighted in Figure 2a (toehold + stem1) kept constant across the library? If not, and if variations in stem1 length are compensated by changes in toehold length, it would be important to clarify whether the authors evaluated the impact of shorter toeholds on switching kinetics and overall performance.
2. Although Detectron optimization improves ON/OFF ratios (notably for variant #621), the absolute performance of the sensing system remains somewhat unclear. Detection of phage infection in pooled cultures was performed at a fixed multiplicity of infection (MOI = 5). It would be important to understand how the system performs at lower MOI. Can the authors define the dynamic range of detection and the sensitivity of the system with respect to infection levels?

3. The authors describe the system as enabling “digital” sensing; however, barcode abundance is quantified by normalization to plasmid levels and expressed as fold changes between infected and non-infected samples. In this context, the use of the term “digital,” referred to a “biosensor,” may be somewhat misleading. In particular, it is not clear how Detectrons compare quantitatively with existing RNA sensing systems (e.g., toehold-based reporters or CRISPR-based diagnostics). Providing a benchmark—either experimentally and through a more explicit discussion—would help contextualize the performance of this platform and clarify its main advantages. In addition, the requirement for sequencing to read out DNA barcodes should be discussed in terms of cost, scalability, and practical implementation of the platform.

4. The level of multiplexing demonstrated experimentally (three phages and a limited number of barcodes) remains relatively modest. The authors should increase the number of infected phages to demonstrate the orthogonality of the detection. In general, it is unclear what the practical limits of the system are in terms of library size, sequencing depth, potential crosstalk, and barcode bias.

5. The phage detection experiments rely on chloramphenicol treatment to delay lysis and allow barcode accumulation. While this is a clever and practical solution, it raises some concerns. In particular, it is unclear to what extent this intervention perturbs the natural infection dynamics or affects transcriptional profiles, which could in turn influence the sensing output. The authors should discuss these aspects and clarify whether the system can function under more physiological conditions, or without such intervention.

6. Some statements in the Discussion, particularly regarding the potential extension of Detectrons to mammalian systems, appear speculative and would benefit from more cautious framing.

This manuscript presents a creative and technically strong platform for RNA sensing via DNA barcode generation, with promising applications. The integration of programmable RNA recognition with retron-based DNA outputs is elegant and potentially impactful.

However, to meet the standard of Nature Biotechnology, the manuscript would benefit from a clearer articulation of its conceptual advance, a more critical positioning relative to existing technologies, and a more careful framing of its claims—particularly regarding its impact in biosensing and its broader applicability.

With these revisions, the work would represent a valuable contribution to the field of synthetic biology and biosensing.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

Han et al. have developed the Detectrons, which couples programmable toehold switches with retron-mediated reverse transcription to transduce RNA inputs into unique DNA barcodes. Through the combination of generating toehold-retron library and machine learning strategy, they uncover the design principles of Detectrons. Then the optimized version of Detectrons were used for specific phage infection detection and can be extended to other bacteria species. Overall, the manuscript is well written and this work presents a novel approach to transduce RNA signal into unique DNA barcodes in cells. However, its biotech applications are very limited to phage-host interaction screening at present, and hence have not been sufficiently demonstrated. In addition, the following points should be addressed.

Major comments :

1. In the introduction, the authors mentioned that “while protein outputs can be useful for real-time and functional readouts, they significantly lack the diversity, sensitivity, programmability, and multiplexing capacity of DNA outputs, which are more stable and better suited for biosensors and synthetic circuits.” However, substantial evidence indicates that intracellular accumulation of single-stranded DNA (ssDNA)—a common form of DNA-based output in bacterial systems—acts as a potent genomic stress signal: it triggers the RecA-mediated SOS response, induces cell-cycle arrest (e.g., via SulA), elevates mutagenesis, and may undergo rapid degradation by nucleases. These well-documented cytotoxic and destabilizing effects appear to directly contradict the claimed “stability” and reliability of DNA outputs *in vivo*. Could the authors clarify: (i) Whether their framework specifically addresses ssDNA toxicity; (ii) How the theoretical advantages of DNA outputs (multiplexing, programmability) are practically balanced against the physiological burden and signal instability caused by ssDNA accumulation in living *E. coli*; and (iii) Under what experimental conditions DNA outputs demonstrably outperform optimized fluorescent protein reporters in terms of sensitivity, dynamic range, or long-term fidelity within cellular environments?

2. The authors developed a high-throughput variant library to optimize the conformational dynamics of Detectrons and to uncover their design principles. In Figure 4, Detectrons-V2 exhibit lower leakiness compared to V1, as seen in the gel image. However, in Figure 5, high leakiness is observed, even for the same phage. This discrepancy raises concerns about whether ssDNA extraction is a reliable quantitative method for evaluating Detectrons, as the extraction efficiency may vary and cannot be guaranteed to be consistent across samples.

3. In Figures 4 and 5, the authors tested only one target site per phage. Since the Toehold switch relies on RNA interactions, and the secondary structure of mRNA can significantly influence interaction efficiency, I suggest that the authors examine additional target sites during phage detection for comparison.

4. Retrons produce msDNA, which is a specific RNA-DNA complex consisting of a single-stranded DNA linked to a small RNA molecule via a covalent bond. In the retron complex, the msDNA complex is associated with RT, and this will affect the

extraction of msDNA from cells. Does the author believe that plasmid extraction is sufficient to remove RT from the msDNA complex, or will the complex be precipitated during extraction? This could significantly affect the quantitative results.

5. I would suggest that the authors link the Detectron with the previously developed molecular recording tool (Nature. 2022 Aug; 608 (7921): 217-225) to generate a stable "output signal" on a plasmid.

(Remarks on code availability)

Author Rebuttal letter:

We appreciate the time and effort dedicated to thoughtfully reviewing our original manuscript. We have carefully revised the work, incorporating new analyses, experiments, clarifications, and discussion points raised by the reviewers, resulting in an improved manuscript. Below is a point-by-point response to each reviewer.

Reviewer #1 (Remarks to the Author):

This manuscript describes a novel and broadly useful "Detectron" that enables conversion of RNA signals into a more durable DNA output signal in bacteria. Detectrons operate through a toehold switch-inspired regulatory mechanism where a 5' hairpin structure, which initially blocks retron reverse transcription, is unwound by a complementary trigger RNA, promoting RNA refolding and activation of retron reverse transcription. The authors demonstrate that Detectrons can be used to detect natural RNAs (e.g. Zika virus fragments, phage transcripts) to induce production of barcoded DNA reporters, providing longer-lived products that can be recorded through downstream sequencing. The manuscript provides a detailed study of Detectron design parameters based on a library of ~7800 analyzed variants and further uses machine learning to identify the most salient system features. They confirm that the system can work in multiple bacterial hosts and demonstrate that phage-specific Detectrons can be used to record the preferences of phage infection across a population of different strains of bacteria.

Overall, the manuscript is very clearly written, the results support the robust operation of the system, and the testing and optimization based on a high-throughput library of Detectrons goes well beyond the norm for the first paper describing a new RNA-based regulatory system. The Detectrons themselves represent a valuable new capability for synthetic biology, which could be deployed for a multitude of different applications where long-term storage of transient RNA information is needed. For publication in a high-impact journal like Nature Biotechnology, the main weakness of the paper is the deployment of Detectrons for phage-host interactions, where experiments are limited to three hosts or phages and it is unclear how much the method improves over more conventional approaches. I recommend publication once the points below are addressed:

1. The authors should provide more details about the plasmid design for the Detectron system. Critical components of the Detectron plasmid are not sufficiently described, such as the origin of plasmid, the RT expression, etc. Those details are essential for readers to understand and reproduce the experiments.

We thank the reviewer for this comment. We have now clarified the plasmid architecture in both the Methods section and Supplementary Table 2. Specifically, we summarized key components of the plasmids used in this study, including promoter, terminator, origin of replication, selection marker, and plasmid backbone:

Detectron constructs were cloned into an ampicillin-resistant pBR322-derived plasmid containing a ColE1/pMB1-family origin together with *bom* and *rop*. The retron ncRNA (*msr-msd*) and Eco1 reverse transcriptase were expressed as a polycistronic cassette under the constitutive *glpT* promoter and terminated using B0014 terminator. Trigger RNAs were encoded on the same plasmid under the indicated promoters (*J23100*, *J23105*, or *pBAD*) and terminated using the T7 transcription terminator (*T7Te*). Detailed construct annotations are provided in Supplementary Table 2.

2. It is unclear in the manuscript how the Detectron approach improves over other approaches that could be used to examine phage-host interactions. For instance, could DNA sequencing of surviving cells or phage populations after infection be used to examine phage-host interactions with similar outcomes? The need to supply chloramphenicol to suppress cell lysis seems

as if it could limit the approach when applied to more diverse bacterial communities for phage-host studies. Providing a clearer and more detailed description of the benefits and limitations of the Detectrons would be valuable for readers.

We thank the reviewer for this important comment. We have revised the manuscript to more clearly describe both the advantages and limitations of the Detectron approach:

Existing sequencing-based approaches for studying phage-host interactions primarily profile surviving cells or phage populations after infection. As a result, these methods capture selection outcomes rather than directly reporting early infection events. While phage transcripts can in principle be detected by RNA sequencing, such approaches measure transient RNA species and are limited by the dominance of host RNA, the loss of signal due to rapid cell lysis, and the lack of inherent linkage between infection signals and specific host-phage pairs in multiplexed settings. In contrast, Detectrons convert RNA inputs into stable DNA barcodes that can be readily recovered and quantified by sequencing. By linking barcode identity to

specific phage sensors and host genotypes, Detectrons enable sequence-resolved analysis of phage–host interactions within a pooled culture, a capability not readily achievable with existing sequencing-based approaches.

While Detectrons enable multiplexed and sequence-resolved analysis of phage–host interactions, it has several limitations. In our current implementation, chloramphenicol was used to suppress phage-induced lysis, allowing sufficient time for barcode accumulation and subsequent recovery. As a result, the system is best suited for comparing phages with similar infection kinetics, as phages with substantially different replication or lysis rates may not be accurately resolved within a single reaction. We note that chloramphenicol treatment may alter aspects of host physiology and transcriptional dynamics during infection. However, because Detectrons record the presence of specific phage transcripts rather than absolute transcriptional output, the system primarily captures relative infection signals within a controlled experimental context, and comparisons across conditions remain internally consistent. Future implementations that decouple barcode generation from host survival, or enable more rapid signal recording, may extend this approach to more diverse phage communities and allow operation under more physiological conditions.

3. The first sentence of the Results section is confusing and should be revised. It currently suggests that the retron ncRNA also contains a single-stranded DNA region that is also reverse transcribed; the timing of the conversion of msd RNA to DNA should be clarified.

We agree that the original wording could be misinterpreted as implying that the msd region is DNA prior to reverse transcription. To clarify this point, we have revised the first sentence as follows:

The retron ncRNA consists of two distinct, partially overlapping regions: the msr (multicopy single-stranded RNA), which is not reverse transcribed, and the msd (multicopy single-stranded DNA–encoding region), which serves as the template for reverse transcription by the retron reverse transcriptase.

4. It would be helpful to have arrowheads indicating the 3' ends of the RNAs/DNAs in the figures. This is particularly true in Fig. 1a where there is the trigger, retron RNA, and nascent ssDNA (which confusingly is also the same color as the a1/a2 region).

We have revised Fig. 1a and 1b to include arrowheads indicating the 3' ends of all RNA and DNA species, improving the clarity of strand directionality. In addition, we changed the color of the nascent ssDNA to clearly distinguish it from the a1/a2 region.

Figure 1. Detectrons specifically report on target RNA by DNA barcode synthesis. a, Schematic of the biological components of a Detectron. RT, reverse transcriptase. b, Schematic of the Detectron mechanism.

5. Did the authors test any Detectrons with loop lengths shorter than 12 nt or are there system features (e.g. a1/a2 lengths) that prevented this? The ML results and RNA folding considerations suggest that shorter loops could lead to more stable structures and lower signal leakage. More information should be provided as to why certain ranges of domain lengths were selected.

We thank the reviewer for this important comment. As suggested, we performed new experiments to test additional Detectrons with loop lengths shorter than 12 nt, including 11, 9, 7, 5, 3, and 1 nt. Because a2 was originally defined as the sequence spanning from the loop, excluding the first three nucleotides, to the priming guanosine, changing the loop length would also alter a2 length and could confound interpretation of RT-DNA production. To isolate the effect of loop length itself, we designed a focused set of constructs in which the loop was excluded from the a2 sequence, thereby fixing a2 length at 12 nt while varying only the loop length (Extended Data Fig. 6a).

In this analysis, shorter loop lengths increased ON-state RT-DNA production, leading to improved ON/OFF ratios (Extended Data Fig. 6b,c). The strongest performance was observed with a 5-nt loop, which produced a 3.8-fold higher ON/OFF ratio than the 12-nt loop. OFF-state RT-DNA production was not substantially reduced, indicating that the improved ON/OFF ratios were primarily driven by enhanced ON-state output rather than reduced leakiness. These results indicate that shorter loops can enhance activated-state RT-DNA production when a2 length is held constant, suggesting that loop length influences the efficiency of the activated conformation rather than simply reducing OFF-state leakage. We have added this analysis to the revised manuscript (Extended Data Fig. 6).

Extended Data Fig. 6 Effect of shorter loop lengths on Detectron performance. a, Schematic of the focused loop-length analysis. To isolate the effect of loop length

from changes in a2 length, a2 was fixed at 12 nt by excluding the loop from the a2 sequence, while loop length was varied from 12 nt to shorter lengths of 11, 9, 7, 5, 3, and 1 nt. b, Enrichment of RT-DNA + plasmid over plasmid alone for Detectrons with the indicated loop lengths, as measured by qPCR; One-way ANOVA with Tukey's multiple comparisons test (corrected) for ON-state signal: 7 nt versus 12 nt: $P < 0.0001$; 5 nt versus 12 nt: $P < 0.0001$; 3 nt versus 12 nt: $P < 0.0001$. Closed circles indicate three biological replicates. Bars represent the mean (s.d.) of three biological replicates. c, ON/OFF ratios of Detectrons with the indicated loop lengths; One-way ANOVA with Tukey's multiple comparisons test (corrected): 7 nt versus 12 nt: $P < 0.0001$; 5 nt versus 12 nt: $P < 0.0001$; 3 nt versus 12 nt: $P < 0.0001$. Bars represent the mean (s.d.) of three biological replicates.

6. Some additional information concerning the Detectron library would be helpful to state in the main text. Was the a1/a2 length varied? Were the stem2 sequences randomized? Did all elements in the library use the same trigger sequence?

The a1/a2 length did vary based on the stem length, a single stem2 sequence was used across all library members, and all elements in the library use the same trigger sequence.

We acknowledge that there is complexity in the design elements, so we have added text, as well as additional diagrams and explicit sequence details in revised Fig. 2a and Extended Data Fig. 3.

The target-binding domain (a + b), comprising the fixed trigger-binding region (a, 25 nt) and the variable stem1 region (b, 7-20 nt), mediates hybridization with a universal trigger RNA, used across the entire library (Fig. 2a).

Figure 2. Optimization of Detectrons using a high-throughput variant library. a, Schematic of the Detectron variant library design. stem1, stem2, bulge, loop, and the distance from stem base to the priming guanosine (d) were systematically varied to construct 7,840 variants. All parameters were varied in 1-nt increments unless otherwise indicated; loop length was varied in 3-nt increments.

The effective length of the interacting region was not fixed but varied implicitly with the structural parameters of each variant. Specifically, a2 was defined as the sequence spanning from the loop (excluding the first three nucleotides) to the priming guanosine, such that its length depended on the loop, stem2, bulge, stem1, and d parameters (Extended Data Fig. 3a). Stem2 sequences were not randomized but deterministically defined for each length (0-6 nt; Extended Data Fig. 3b). All variants were evaluated using a single, fixed trigger RNA (135 nt). This trigger comprises two functional segments: a 25-nt region (a*) complementary to the trigger-binding region (a) and an additional region (b*) that pairs with the variable stem1 region (b) (Fig. 2a). The effective pairing region (a* + b*) therefore varies only with the length of b.

Extended Data Fig. 3 Definition of variable structural elements in the Detectron variant library. a, Schematic of the Detectron switch RNA architecture used for library construction. The target-binding domain consists of the fixed trigger-binding region (a) and the variable stem1 region (b). The a2 region was defined as the sequence spanning from the loop, excluding the first three nucleotides, to the priming guanosine. Consequently, a2 length varied as a function of loop, stem2, bulge, stem1, and d lengths. b, Maximum-length sequences used for each variable region during library design. The trigger-binding region (a) was fixed at 25 nt, whereas stem1 (b), bulge, stem2, loop, and d were varied according to the specified design parameters. Stem2 sequences were deterministically defined such that shorter stem2 variants were generated by progressive truncation of the maximum-length stem2 sequence (ACUUA). Complementary regions (*stem1 and *stem2) were defined as the reverse complements of stem1 and stem2, respectively.

7. A Detectron where the trigger invades into stem2 would probably perform better since it would discourage reversal of the strand-displacement reaction and thus more stable a1/a2 binding upon activation. Were any systems in the library designed so that the trigger invaded into stem2? If not, it may be interesting to revisit the library testing results to see if partial stem2 strand invasion (e.g. 1-3+ base overlap with the trigger) led to any performance improvements.

We thank the reviewer for this insightful suggestion. In our design, stem1 is complementary to the target RNA, whereas stem2 consists of fixed nucleotides. Importantly, the library included constructs lacking stem2 (stem2 = 0), in which the full stem is formed by the trigger-complementary stem1 region. These constructs therefore allow us to test whether extending

trigger-complementary pairing into the portion of the stem otherwise occupied by stem2 improves performance. We have included this new analysis in the results:

To further assess whether extending the trigger-complementary region improves switching behavior, we analyzed a subset of variants in which bulge, loop, and d lengths were held constant (bulge = 0, loop = 12, d = 1) and stem2 was absent (stem2 = 0), such that the full stem consisted of the trigger-complementary stem1 region (Extended Data Fig. 9a). ON/OFF ratios peaked at stem1 lengths of approximately 9–11 bp and declined with further increases in stem1 length (Extended Data Fig. 9b). These results indicate that extending trigger-complementary pairing along the stem does not monotonically improve performance but instead exhibits an optimal range.

Extended Data Fig. 9 Effect of extended trigger-complementary stem pairing on Detectron performance. a, Schematic of variants used to test extended trigger-complementary pairing along the stem. Bulge, loop, and d lengths were held constant at 0 nt, 12 nt, and 1 nt, respectively, and stem2 was absent (stem2 = 0), and stem1 length was varied such that the full stem consisted of the trigger-complementary stem1 region. b, Normalized ON/OFF ratios of Detectron variants with the indicated stem1 lengths.

8. Toehold-mediated RNA refolding and stem-formation mechanisms similar to those used in the Detectrons have been used in a few papers (e.g. PMID: 39127751, 39036067, 41272150). These earlier papers should be cited.

We agree. We have now cited these studies in the Introduction in a sentence describing prior applications of toehold-mediated strand displacement and RNA refolding mechanisms in synthetic RNA systems.

9. For Fig. 1h, the signal from cells with Detectron alone is higher than Detectron + scrambled RNA, which is unexpected and perhaps a statistically significant difference. Is the increased signal due to the scrambled RNA interacting with the Detectron to inhibit its activation?

We thank the reviewer for this comment. Although the Detectron + scrambled RNA condition appears slightly lower than Detectron alone, statistical analysis showed that this difference was not significant (one-way ANOVA with Tukey's multiple comparisons test, adjusted P = 0.4533; Supplementary Table 1). Therefore, our data do not support the interpretation that scrambled RNA inhibits Detectron activation.

10. The ON/OFF ratio for T7-V2 exceeds 30 in Figure 4 but is reduced to 2 in Figure 5c. Please explain why this large drop occurs.

We thank the reviewer for this insightful comment. The values reported in Fig. 4 and Fig. 5c are derived from different experimental contexts and are therefore not directly comparable. Fig. 4 reports ON/OFF ratios measured using plasmid-expressed target mRNAs in a controlled two-state system, whereas Fig. 5c quantifies barcode enrichment following native phage infection in a mixed culture. In this sequencing-based infection assay, barcode abundance is normalized to the wt barcode from the No Toehold control strain, and signal accumulation occurs within a limited time window prior to phage-induced lysis. Together with differences in target mRNA abundance during infection, these factors compress the apparent dynamic range relative to the ON/OFF measurements in Fig. 4. We have clarified this point in the revised Results section.

It is important to note that the dynamic range observed in this sequencing-based infection assay is not directly comparable to the ON/OFF ratios measured using plasmid-expressed target mRNAs in Fig. 4. In contrast to the controlled ON/OFF comparison, this assay quantifies barcode enrichment following native phage infection in a mixed culture and incorporates normalization to the wt barcode from the No Toehold control strain. Together with differences in trigger RNA abundance and the limited barcode accumulation window during infection, these factors compress the apparent dynamic range, resulting in reduced ON-state signal relative to OFF and an apparent increase in baseline signal despite similar underlying leakiness. Importantly, normalization to the constitutively produced wt barcode from a separate positive control strain provides a pooled reference for comparing infection-dependent barcode enrichment across samples.

11. For the machine learning model, the authors should provide more details in how they define the SHAP baseline for both feature importance score and ON/OFF ratio across different feature distributions.

We thank the reviewer for this helpful comment. We have clarified the SHAP analysis in the Methods. Briefly, after selecting the best-performing feature set and hyperparameters by nested cross-validation, we retrained a final XGBoost model on the full dataset and computed SHAP values for all variants using this final model. For Fig. 3c, feature importance scores were defined as the mean absolute SHAP value for each feature across all variants. For Fig. 3d–h, the SHAP values shown for each feature correspond to per-variant contributions relative to the same baseline, and bars represent the mean SHAP value

for variants sharing the same feature value or feature bin.

SHAP values were computed using the full standardized feature matrix as the background dataset, such that the SHAP baseline corresponds to the expected model prediction across all variants. Individual SHAP values therefore represent the contribution of each feature to shifting the predicted ON/OFF ratio from this baseline for a given variant, and mean SHAP values were calculated across variants sharing the same feature value or feature bin to generate the distributions shown in Fig. 3d-h.

12. In Figure 3b, only the test set data are shown, but the caption mentions the training and test set data being present. Please revise.

We appreciate the reviewer's comment. We have revised the caption to clarify that Fig. 3b displays only the test set data, removing the reference to training set data.

13. Page 16: "onT7" should be "on T7".

We thank the reviewer for pointing this out. We have corrected "onT7" to "on T7" on page 16.

14. In Extended Data Fig. 1, the priming G site should be located closer to the 5' end rather than the 3' end in the ECO1 ncRNA.

We thank the reviewer for pointing this out. We have revised Extended Data Fig. 1 to position the priming G site closer to the 5' end of the Eco1 ncRNA.

Reviewer #2 (Remarks to the Author):

This manuscript by Han and Shipman introduces "Detectrons," a synthetic biology platform that couples programmable RNA sensing via toehold switches with retron-mediated DNA barcode generation. By converting transient RNA signals into stable DNA outputs, the authors propose a new framework for multiplexed sensing and recording of transcriptional events. The work is technically sophisticated, combining circuit design, high-throughput screening, and machine learning, and is applied to an interesting use case, namely phage-host interaction profiling in pooled bacterial cultures.

The authors position Detectrons as the first RNA-to-DNA transduction system. While the implementation is elegant, the conceptual distinction from prior molecular recording systems (e.g., retron-based recording or CRISPR-based acquisition systems) is not fully articulated, particularly in the Introduction. The manuscript would benefit from a clearer and more explicit comparison with recent retron-based recording approaches and RNA-guided recording strategies. As it stands, the advance appears at least in part incremental—i.e., based on the integration of existing components (toehold switches and retrons)—rather than entirely conceptual. It would therefore be important for the authors to better define what fundamentally new capability is enabled by Detectrons and in which aspects this system clearly outperforms existing approaches.

Overall, the manuscript is well organized and the experimental workflow is clearly described. However, several aspects require clarification and, in some cases, additional support to fully justify the scope of the claims and strengthen the overall impact.

Major comments

1. The hypothesis that sequestration of the $\alpha 2$ region inhibits reverse transcription and that strand displacement restores activity is supported by the experimental data, although no direct structural or biochemical validation is provided. Given that the observed improvements in ON/OFF ratios are attributed to structural parameters (stem1, stem2, loop, and bulge), it is unclear whether the effect of toehold length itself was systematically investigated. In particular, is the length of the region highlighted in Figure 2a (toehold + stem1) kept constant across the library? If not, and if variations in stem1 length are compensated by changes in toehold length, it would be important to clarify whether the authors evaluated the impact of shorter toeholds on switching kinetics and overall performance.

We thank the reviewer for this insightful comment regarding the potential impact of toehold length on switching performance.

In our library design, the toehold region (a) was fixed at 25 nt across all variants, whereas only the stem1 region (b) was varied (7–20 nt), as shown in Fig 2a. Accordingly, the total length of the target-binding domain (a + b) was not constant and varied solely due to changes in stem1, and not due to variation in the toehold itself. This design allowed us to isolate the effects of structural parameters (stem1, stem2, loop, and bulge) without confounding variation in toehold length.

Consistent with this design, all variants were evaluated using a single, fixed trigger RNA (135 nt), in which the region complementary to (a + b) varied only in length according to stem1. Thus, the toehold length remained constant across the entire library, and we did not systematically evaluate the effects of shorter toeholds on switching kinetics or performance in this study. We have clarified this design explicitly in the revised manuscript:

All variants were evaluated using a single, fixed trigger RNA (135 nt). This trigger comprises two functional segments: a 25-nt region (a*) complementary to the trigger-binding region (a) and an additional region (b*)

that pairs with the variable stem1 region (b) (Fig. 2a). The effective pairing region ($a^* + b^*$) therefore varies only with the length of b.

Figure 2. Optimization of Detectrons using a high-throughput variant library. a, Schematic of the Detectron variant library design. stem1, stem2, bulge, loop, and the distance from stem base to the priming guanosine (d) were systematically varied to construct 7,840 variants. All parameters were varied in 1-nt increments unless otherwise indicated; loop length was varied in 3-nt increments.

2. Although Detectron optimization improves ON/OFF ratios (notably for variant #621), the absolute performance of the sensing system remains somewhat unclear. Detection of phage infection in pooled cultures was performed at a fixed multiplicity of infection (MOI = 5). It would be important to understand how the system performs at lower MOI. Can the authors define the dynamic range of detection and the sensitivity of the system with respect to infection levels?

We thank the reviewer for this suggestion. To define the dynamic response of Detectron across infection levels, we performed an additional T7 phage infection experiment across a range of MOIs (0.5, 1, 2, 5, and 10). Barcode enrichment increased with increasing MOI, demonstrating that Detectron output responds quantitatively to infection level (Extended Data Fig. 12). Under the conditions tested, barcode enrichment became apparent at an MOI of 5 and increased further at MOI of 10. These results define the dynamic response of the T7 Detectron and provide an estimate of its sensitivity to phage infection level. We have added these data to the revised manuscript.

Extended Data Fig. 12 MOI-dependent response of T7-targeting Detectron. Barcode enrichment from the T7-targeting Detectron strain following infection with T7 phage at the indicated multiplicities of infection (MOIs). Barcode abundance was normalized to the wt barcode and plotted as fold change relative to the no-infection control. Bars represent the mean (s.d.) of three biological replicates.

3. The authors describe the system as enabling “digital” sensing; however, barcode abundance is quantified by normalization to plasmid levels and expressed as fold changes between infected and non-infected samples. In this context, the use of the term “digital,” referred to a “biosensor,” may be somewhat misleading. In particular, it is not clear how Detectrons compare quantitatively with existing RNA sensing systems (e.g., toehold-based reporters or CRISPR-based diagnostics). Providing a benchmark—either experimentally and through a more explicit discussion—would help contextualize the performance of this platform and clarify its main advantages. In addition, the requirement for sequencing to read out DNA barcodes should be discussed in terms of cost, scalability, and practical implementation of the platform.

We thank the reviewer for this thoughtful comment. We have clarified these points in the revised manuscript.

First, we agree that the term “digital” may be ambiguous in this context. We have therefore revised the text to emphasize that Detectron produces a sequence-encoded output, in which each variant generates a unique DNA barcode that can be identified and quantified by sequencing. This readout is not strictly binary, but instead enables quantitative measurements within pooled samples.

Second, we have expanded the Discussion to better contextualize Detectron relative to existing RNA sensing approaches. While toehold-based reporters and CRISPR-based diagnostics are optimized for sensitive detection of individual targets using fluorescence or cleavage-based readouts, Detectron instead produces DNA barcodes that enable sequence-resolved analysis of multiple targets or conditions within a single experiment. As such, Detectron is complementary to these systems and is particularly suited for multiplexed measurements rather than single-target detection:

Unlike previous toehold switches, which translate RNA inputs into protein-based readouts such as fluorescence reporters⁹, or CRISPR-based detection systems, which rely on programmable nucleic acid cleavage and collateral reporter activation^{27–29}, or those that use engineered crRNAs or guide RNAs to indirectly sense RNA expression via DNA targets³⁰, Detectrons directly generate a DNA signal. Detectron differs fundamentally from these systems in that it produces sequence-encoded DNA barcodes that can be pooled and quantified by sequencing, enabling scalable and multiplexed measurements across many targets or conditions in parallel. As such, Detectron is not designed to maximize single-target sensitivity, but rather to enable high-throughput, sequence-resolved analysis of RNA-dependent processes through DNA barcode readout. This DNA-based output eliminates reliance on protein or fluorescence-based reporters and provides a stable molecular signal. Here, “sequence-encoded” refers to the ability of distinct DNA barcode sequences to represent specific RNA inputs, host genotypes, or experimental conditions, allowing signals to be discretely counted and resolved by sequencing.

Finally, we have added discussion regarding the use of sequencing as a readout. While sequencing introduces considerations of cost and turnaround time, it also enables pooled analysis of multiple barcodes within a single experiment. This capability supports parallel measurements across conditions and makes Detectron well suited for applications where relative comparison across multiple targets or conditions are of interest.

Sequencing-based readout introduces practical considerations, including cost and turnaround time. However, it also enables multiple DNA barcodes to be quantified within a single experiment, allowing sequence-resolved comparisons across pooled samples or conditions. In principle, barcode capacity can be expanded by increasing barcode length while maintaining sufficient Hamming distance among barcode sequences to reduce misassignment from sequencing or amplification errors. In practice, multiplexing is constrained by construct generation, transformation or library coverage, unequal barcode production, PCR amplification bias, sensor crosstalk, and biological differences in target transcript abundance or infection kinetics. The ratiometric Detectron architecture introduced here mitigates several quantitative biases by normalizing each infection-responsive dual barcode to a strain-matched constitutive reference barcode, but it does not eliminate the need to validate sensor orthogonality and barcode performance for each multiplexed application. Thus, Detectron is particularly well suited for multiplexed relative measurements rather than real-time single-target readout.

4. The level of multiplexing demonstrated experimentally (three phages and a limited number of barcodes) remains relatively modest. The authors should increase the number of infected phages to demonstrate the orthogonality of the detection. In general, it is unclear what the practical limits of the system are in terms of library size, sequencing depth, potential crosstalk, and barcode bias.

We thank the reviewer for raising this important point. In the revised manuscript, we have expanded the experimental demonstration of multiplexing and clarified the practical limits of the system. Rather than increasing the number of phage species tested, we chose to expand the host-susceptibility assay to a pooled panel of five *E. coli* LPS core biosynthesis mutants together with WT, to illustrate how closely related host genotypes can be distinguished within a single pooled assay (Fig. 6d). This revised experiment now demonstrates that Detectrons can resolve infection-associated barcode outputs across a larger set of genetically distinct host strains in a single pooled assay.

To improve quantitative robustness in this expanded multiplexed setting, we also redesigned the host-susceptibility assay using a ratiometric Detectron architecture (Fig. 6c). In this system, each detecting strain produces two barcode outputs: an infection-responsive dual-information barcode encoding the targeted phage and host genotype, and a constitutive strain-matched reference barcode. The responsive/reference ratio provides internal normalization for each host genotype, reducing the impact of strain abundance differences, DNA recovery variation, sequencing-depth variation, and barcode-specific bias. Using this approach, we quantified relative T7 susceptibility across six host genotypes in a single pooled culture and observed a graded response consistent with the expected extent of LPS core alteration (Fig. 6f):

The ratiometric barcode readout revealed a graded T7-triggered response across the LPS core mutant panel (Fig. 6f). WT and mildly altered LPS mutants ($\Delta waaQ$ and $\Delta waaB$) showed fold-change values above the no-infection baseline, indicating a T7-induced increase in the dual/reference barcode ratio. The moderately altered $\Delta waaJ$ strain showed an intermediate response, with a mean fold-change value slightly above the no-infection baseline, suggesting that moderate LPS alteration still permits detectable T7-triggered ratiometric activation, although this response was significantly lower than those observed in WT and the mildly altered mutants. In contrast, severely altered LPS mutants ($\Delta waaC$ and $\Delta waaE$) showed responses at or below the no-infection baseline and were significantly reduced compared with WT, mildly altered mutants, and $\Delta waaJ$, indicating that severe LPS core truncation strongly impairs T7-triggered ratiometric activation. Together, these findings demonstrate that ratiometric Detectrons can resolve relative host susceptibility to a defined phage challenge across multiple genotypes within a single pooled assay.

Figure 6. Ratiometric Detectrons enable multiplexed assessment of host susceptibility to phage infection. a, Schematic of Detectron-V2 devices constructed using Cxx1, Eas1, and Ebu1 retron systems, each with its cognate RT. b, Urea PAGE analysis of RT-DNA production from Cxx1-, Eas1-, and Ebu1-based Detectron-V2 devices in the absence (lane 1, 3, 5) or presence (lane 2, 4, 6) of trigger RNA. Arrowheads indicate the expected RT-DNA products from each retron system. c, Schematic of the ratiometric Detectron-V2-Ebu1 architecture. Each detecting strain contains a phage-targeting Detectron-V2-Ebu1 that produces an infection-responsive dual-information barcode encoding phage and host-strain identity, together with a No-Toehold Ebu1 reference retron that constitutively produces a strain-specific reference barcode. d, Barcode mapping and predicted LPS core structures of the *E. coli* host panel used for T7 susceptibility profiling. Mutants were grouped by expected extent of LPS core alteration. e, Schematic of single-pot T7 susceptibility profiling using pooled detecting strains. f, Quantification of ratiometric barcode induction by sequencing. Fold-change in infection-responsive dual barcode abundance relative to the strain-matched reference barcode is shown relative to the no-infection control; One-way repeated-measures ANOVA with Geisser-Greenhouse correction, followed by Tukey's multiple comparisons test (corrected): WT versus $\Delta waaQ$: $P = 0.0003$; WT versus $\Delta waaB$: $P = 0.0408$; WT versus $\Delta waaJ$: $P = 0.007$; WT versus $\Delta waaC$: $P = 0.0047$; WT versus $\Delta waaE$: $P = 0.0052$; $\Delta waaQ$ versus

$\Delta waaJ$: $P < 0.0001$; $\Delta waaQ$ versus $\Delta waaC$: $P < 0.0001$; $\Delta waaQ$ versus $\Delta waaE$: $P = 0.0001$; $\Delta waaB$ versus $\Delta waaJ$: $P = 0.0002$; $\Delta waaB$ versus $\Delta waaC$: $P < 0.0001$; $\Delta waaB$ versus $\Delta waaE$: $P < 0.0001$; $\Delta waaJ$ versus $\Delta waaC$: $P = 0.0128$; $\Delta waaJ$ versus $\Delta waaE$: $P = 0.0413$. Closed circles indicate nine biological replicates. Bars represent the mean of nine biological replicates.

We have also expanded the Discussion to address practical scaling considerations:

In principle, barcode capacity can be expanded by increasing barcode length while maintaining sufficient Hamming distance among barcode sequences to reduce misassignment from sequencing or amplification errors. In practice, multiplexing is constrained by construct generation, transformation or library coverage, unequal barcode production, PCR amplification bias, sensor crosstalk, and biological differences in target transcript abundance or infection kinetics. The ratiometric Detectron architecture introduced here mitigates several quantitative biases by normalizing each infection-responsive dual barcode to a strain-matched constitutive reference barcode, but it does not eliminate the need to validate sensor orthogonality and barcode performance for each multiplexed application. Thus, Detectron is particularly well suited for multiplexed relative measurements rather than real-time single-target readout.

5. The phage detection experiments rely on chloramphenicol treatment to delay lysis and allow barcode accumulation. While this is a clever and practical solution, it raises some concerns. In particular, it is unclear to what extent this intervention perturbs the natural infection dynamics or affects transcriptional profiles, which could in turn influence the sensing output. The authors should discuss these aspects and clarify whether the system can function under more physiological conditions, or without such intervention.

We thank the reviewer for this important point. We agree that the use of chloramphenicol to delay lysis introduces a non-physiological condition that may influence aspects of host physiology and transcriptional dynamics during infection. In our current implementation, chloramphenicol was used to extend the time window for barcode accumulation prior to cell lysis, enabling robust recovery and quantification of Detectron-derived DNA signals.

Importantly, Detectron is designed to report the presence of specific phage-derived transcripts, rather than absolute transcriptional output. As a result, the system primarily captures relative infection signals within a controlled experimental context, and comparisons across conditions remain internally consistent despite potential global perturbations to transcription.

We acknowledge that differences in phage infection kinetics and the effects of translation inhibition may influence the quantitative dynamics of barcode accumulation. However, in the current implementation, Detectron primarily reports early transcriptional events following infection, as barcode generation is triggered by the presence of phage-derived transcripts shortly after entry. Consistent with this, our single-phage infection experiments across multiple bacterial strains demonstrate that Detectron can reliably distinguish infection by a given phage, reflecting successful entry and early gene expression rather than later-stage infection dynamics. Accordingly, in the revised manuscript we clarify that the current implementation is best suited for comparing phages with similar infection kinetics under standardized conditions:

While Detectrons enable multiplexed and sequence-resolved analysis of phage–host interactions, it has several limitations. In our current implementation, chloramphenicol was used to suppress phage-induced lysis, allowing sufficient time for barcode accumulation and subsequent recovery. As a result, the system is best suited for comparing phages with similar infection kinetics, as phages with substantially different replication or lysis rates may not be accurately resolved within a single reaction. We note that chloramphenicol treatment may alter aspects of host physiology and transcriptional dynamics during infection. However, because Detectrons record the presence of specific phage transcripts rather than absolute transcriptional output, the system primarily captures relative infection signals within a controlled experimental context, and comparisons across conditions remain internally consistent.

We agree that extending Detectron to operate under more physiological conditions is an important future direction. Approaches that decouple barcode generation from host survival, or enable more rapid signal recording, may reduce reliance on translation inhibitors and allow application to a broader range of phage–host systems. We have clarified these points in the revised manuscript:

Future implementations that decouple barcode generation from host survival, or enable more rapid signal recording, may extend this approach to more diverse phage communities and allow operation under more physiological conditions.

6. Some statements in the Discussion, particularly regarding the potential extension of Detectrons to mammalian systems, appear speculative and would benefit from more cautious framing.

We thank the reviewer for this comment. We agree that the discussion of potential extensions to mammalian systems was overly speculative in its original form. In the revised manuscript, we have moderated this section to more clearly distinguish

between demonstrated capabilities and future possibilities, and we now explicitly note that substantial optimization would be required to adapt Detectrons to eukaryotic contexts.

This raises the exciting possibility that Detectrons could be adapted to eukaryotic contexts to detect endogenous transcripts generate sequence-encoded DNA barcodes. However, substantial optimization would likely be required to ensure efficient expression, RNA sensing, and DNA production in these settings. This manuscript presents a creative and technically strong platform for RNA sensing via DNA barcode generation, with promising applications. The integration of programmable RNA recognition with retron-based DNA outputs is elegant and potentially impactful.

However, to meet the standard of Nature Biotechnology, the manuscript would benefit from a clearer articulation of its conceptual advance, a more critical positioning relative to existing technologies, and a more careful framing of its claims—particularly regarding its impact in biosensing and its broader applicability.

With these revisions, the work would represent a valuable contribution to the field of synthetic biology and biosensing.

Reviewer #3 (Remarks to the Author):

Han et al. have developed the Detectrons, which couples programmable toehold switches with retron-mediated reverse transcription to transduce RNA inputs into unique DNA barcodes. Through the combination of generating toehold-retron library and machine learning strategy, they uncover the design principles of Detectrons. Then the optimized version of Detectrons were used for specific phage infection detection and can be extended to other bacteria species. Overall, the manuscript is well written and this work presents a novel approach to transduce RNA signal into unique DNA barcodes in cells. However, its biotech applications are very limited to phage-host interaction screening at present, and hence have not been sufficiently demonstrated. In addition, the following points should be addressed.

Major comments :

1. In the introduction, the authors mentioned that “while protein outputs can be useful for real-time and functional readouts, they significantly lack the diversity, sensitivity, programmability, and multiplexing capacity of DNA outputs, which are more stable and better suited for biosensors and synthetic circuits.” However, substantial evidence indicates that intracellular accumulation of single-stranded DNA (ssDNA)—a common form of DNA-based output in bacterial systems—acts as a potent genomic stress signal: it triggers the RecA-mediated SOS response, induces cell-cycle arrest (e.g., via SulA), elevates mutagenesis, and may undergo rapid degradation by nucleases. These well-documented cytotoxic and destabilizing effects appear to directly contradict the claimed “stability” and reliability of DNA outputs in vivo. Could the authors clarify: (i) Whether their framework specifically addresses ssDNA toxicity; (ii) How the theoretical advantages of DNA outputs (multiplexing, programmability) are practically balanced against the physiological burden and signal instability caused by ssDNA accumulation in living *E. coli*; and (iii) Under what experimental conditions DNA outputs demonstrably outperform optimized fluorescent protein reporters in terms of sensitivity, dynamic range, or long-term fidelity within cellular environments?

We thank the reviewer for this insightful comment regarding the potential physiological impact of ssDNA accumulation and the need to more carefully define the practical advantages and limitations of DNA-based outputs.

With respect to ssDNA toxicity and the tradeoff between multiplexing capacity and physiological burden, we agree that intracellular ssDNA can, under certain conditions, trigger host stress responses such as RecA-mediated SOS activation, and that excessive accumulation may impose a physiological burden. We have therefore revised the manuscript to avoid implying that DNA outputs are universally stable or physiologically inert. At the same time, natural retrons have been reported to produce hundreds of msDNA in bacterial cells (Yee et al., *Cell*, 1984), indicating that retron-derived ssDNA can be tolerated within certain biological contexts. Because absolute barcode copy number and host-cell burden were not directly quantified in this study, we now frame expression level and host fitness as important parameters for future optimization. Thus, in the context of Detectron, the advantage of DNA output lies in information persistence and sequence recoverability relative to transient RNA or protein outputs, rather than complete physiological inertness. We have clarified these points in the revised manuscript:

“While DNA-based outputs offer advantages in programmability and multiplexing, we note that intracellular accumulation of ssDNA can, under certain conditions, engage host stress responses such as RecA-mediated SOS pathway. However, natural retrons have been reported to produce hundreds of msDNA in bacterial cells¹⁷, indicating that retron-derived ssDNA can be tolerated within certain biological contexts. Thus, in the context of Detectron, the advantage of DNA output lies in information persistence and sequence recoverability relative to transient RNA or protein outputs, rather than complete physiological inertness. Future implementations should further define the expression regimes that balance barcode production with host-cell fitness.”

Finally, with respect to the comparison between DNA outputs and fluorescent protein reporters, we have revised the manuscript to avoid implying that DNA outputs are universally superior. Protein-based reporters are well suited for real-

time and single-cell measurements, including applications that require optimized sensitivity or dynamic range. On the other hand, Detectrons enable sequence-resolved, multiplexed, and pooled analysis across many targets or conditions in parallel. As such, the advantages of Detectron lie in scalability and multiplexing rather than maximal single-target sensitivity or real-time readout. We have also clarified these points in the revised manuscript:

Detectrons differ fundamentally from these systems in that they produce sequence-encoded DNA barcodes that can be pooled and quantified by sequencing, enabling scalable and multiplexed measurements across many targets or conditions in parallel. As such, Detectrons are not designed to maximize single-target sensitivity, but rather to enable high-throughput, sequence-resolved analysis of RNA-dependent processes through DNA barcode readout.

2. The authors developed a high-throughput variant library to optimize the conformational dynamics of Detectrons and to uncover their design principles. In Figure 4, Detectrons-V2 exhibit lower leakiness compared to V1, as seen in the gel image.

However, in Figure 5, high leakiness is observed, even for the same phage. This discrepancy raises concerns about whether ssDNA extraction is a reliable quantitative method for evaluating Detectrons, as the extraction efficiency may vary and cannot be guaranteed to be consistent across samples.

We thank the reviewer for this thoughtful comment. The apparent difference in leakiness between Fig. 4 and Fig. 5 arises primarily from differences in how the data are visualized, rather than variability in ssDNA extraction.

First, we note that the PAGE gel images in Fig. 4 and Fig. 5 are qualitative and are not intended for direct quantitative comparison across experiments. Differences in overall signal intensity and imaging exposure can substantially affect the apparent band intensity. In particular, because overall signal levels are lower in Fig. 5, higher exposure during imaging can make OFF-state bands appear more prominent relative to ON-state bands, giving the impression of increased leakiness.

In addition, the two experiments are performed in fundamentally different biological contexts. In Fig. 4, Detectron performance was evaluated using plasmid-driven expression of full-length target transcripts under a strong constitutive promoter (J23100), resulting in high and sustained target RNA abundance. In contrast, Fig. 5 reflects detection under native phage infection conditions, where target transcript abundance is transient and barcode accumulation is constrained by a limited time window prior to cell lysis. These factors reduce the absolute ON-state signal in Fig. 5 and compress the apparent dynamic range.

Importantly, the sequencing-based quantification in Fig. 5 relies on normalization to a constitutive wt barcode control within the same sample, enabling robust relative comparisons that are not sensitive to variations in extraction efficiency. Consistent with this, barcode enrichment remains specific to cognate phage infection with minimal cross-activation, supporting the reliability of the assay.

We have clarified these distinctions in the revised manuscript:

It is important to note that the dynamic range observed in this sequencing-based infection assay is not directly comparable to the ON/OFF ratios measured using plasmid-expressed target mRNAs in Fig. 4. In contrast to the controlled ON/OFF comparison, this assay quantifies barcode enrichment following native phage infection in a mixed culture and incorporates normalization to the wt barcode from the No Toehold control strain. Together with differences in trigger RNA abundance and the limited barcode accumulation window during infection, these factors compress the apparent dynamic range, resulting in reduced ON-state signal relative to OFF and an apparent increase in baseline signal despite similar underlying leakiness. Importantly, normalization to the constitutively produced wt barcode from a separate positive control strain provides a pooled reference for comparing infection-dependent barcode enrichment across samples.

3. In Figures 4 and 5, the authors tested only one target site per phage. Since the Toehold switch relies on RNA interactions, and the secondary structure of mRNA can significantly influence interaction efficiency, I suggest that the authors examine additional target sites during phage detection for comparison.

We thank the reviewer for this important suggestion. We agree that the secondary structure and accessibility of the target mRNA can influence RNA-RNA interaction efficiency and therefore affect Detectron activation. To directly compare target-site effects under controlled expression conditions, we generated additional Detectrons targeting four distinct 33-nt regions within the T4 gp23 mRNA and measured their ON/OFF ratios using gp23 mRNA expressed from a strong promoter (Extended Data Fig. 10a).

We also measured the predicted accessibility of each target site using RNAplfold. Because Detectron binding is expected to initiate through the 25-nt region of the target site that hybridizes with the exposed toehold region of the sensor, we calculated the predicted opening energy of this 25-nt target region within the full gp23 mRNA context (Extended Data Fig. 10a). RNAplfold was run using a local folding window of 150 nt, a maximum base-pair span of 70 nt, and unpaired stretch reporting up to 33 nt. Opening energy was calculated as $\Delta G_{\text{open}} = -RT \ln(pU)$, where pU is the predicted probability that the target region is unpaired.

Among the four tested gp23 target sites, only the site 1-targeting Detectron exhibited meaningful ON/OFF ratio, whereas the other three Detectrons showed minimal activation (Extended Data Fig. 10b). Site 1 also had the lowest predicted target-site opening energy, consistent with greater local accessibility contributing to Detectron activation. These results support the reviewer's point that target-site selection and mRNA secondary structure are important considerations for Detectron design.

Although the functional site had the lowest predicted target-site opening energy, we do not infer that target accessibility alone fully determines Detectron performance. Because each target site dictates the complementary sequence incorporated into the Detectron, changing the target region may also alter the folding or activation energetics of the engineered switch RNA. Thus, these results suggest that target-site accessibility is a useful design metric, while empirical testing of multiple candidate sites remains important for identifying functional Detectrons. We have added these new data and revised the manuscript accordingly:

Because Detectron activation depends on RNA–RNA hybridization between the target mRNA and the engineered switch RNA, we next asked whether performance depends on target-site choice within a phage transcript. We designed additional Detectron-V2 devices targeting four distinct 33-nt regions within the T4 gp23 mRNA and measured their ON/OFF ratios using gp23 mRNA (Extended Fig. 10a). To estimate local target-site accessibility, we used RNAplfold to calculate the predicted opening energy (ΔG_{open}) of the 25-nt target region expected to mediate initial Detectron binding within each 33-nt site. Among the four sites, site 1 had the lowest predicted target-site opening energy, indicating the greatest predicted accessibility (Extended Fig. 10a). Consistent with this analysis, only the site 1-targeting Detectron produced a robust ON/OFF ratio, whereas Detectrons targeting the other three sites showed minimal activation (Extended Fig. 10b). These results support the importance of target-site selection and are consistent with local mRNA accessibility contributing to Detectron activation. However, because changing the target site also changes the complementary sequence incorporated into the Detectron, additional sequence-dependent effects on switch folding or activation may also contribute to performance differences. Thus, empirical testing of multiple candidate sites remains important for identifying functional Detectrons.

Extended Data Fig. 10 Target-site selection affects Detectron-V2 performance within gp23 mRNA. a, Schematic of four 33-nt target sites selected within the full-length T4 gp23 mRNA. The 25-nt binding region within each target site represents the sequence expected to mediate initial hybridization with the Detectron-V2 switch RNA. Predicted target-site opening energies (ΔG_{open}) for the 25-nt binding regions were calculated using RNAplfold; lower ΔG_{open} values indicate greater predicted local accessibility b, ON/OFF ratios of Detectron-V2 devices targeting the indicated sites, plotted against the predicted opening energy of the corresponding 25-nt binding region. Closed circles indicate three biological replicates. Bars represent the mean of three biological replicates.

4. Retrons produce msDNA, which is a specific RNA–DNA complex consisting of a single-stranded DNA linked to a small RNA molecule via a covalent bond. In the retron complex, the msDNA complex is associated with RT, and this will affect the extraction of msDNA from cells. Does the author believe that plasmid extraction is sufficient to remove RT from the msDNA complex, or will the complex be precipitated during extraction? This could significantly affect the quantitative results.

We thank the reviewer for this insightful comment regarding the potential association of retron reverse transcriptase (RT) with msDNA and its impact on extraction efficiency. We note that while msDNA consists of a covalently linked RNA–DNA hybrid, the RT itself is associated with the complex through noncovalent interactions.

Our extraction workflow involves alkaline lysis followed by silica-based purification, which are strongly denaturing conditions that disrupt protein–nucleic acids interactions and efficiently remove associated proteins during clarification and wash steps. In addition, the subsequent ssDNA cleanup step further enriches nucleic acids and depletes residual proteins. Based on these considerations, we do not expect RT to remain quantitatively associated with msDNA throughout purification or to significantly affect recovery.

Importantly, our analyses rely on relative quantification in which barcode abundance is normalized to the corresponding plasmid abundance from matched samples. As such, any residual extraction bias would need to differ systematically across conditions to affect the observed ON/OFF ratios. The consistency of barcode measurements across replicates supports the robustness of this approach.

We have clarified these points in the revised manuscript:

To analyze RT–DNA on a PAGE gel, 20 ml of culture was pelleted, and RT–DNAs were extracted using the QIAGEN Midiprep Plasmid Plus Kit. Because retron reverse transcriptase is noncovalently associated with msDNA, alkaline lysis and silica-based purification are expected to disrupt RT–nucleic acid interactions and recover msDNA primarily as nucleic acid species.

5. I would suggest that the authors link the Detectron with the previously developed molecular recording tool (Nature. 2022 Aug; 608 (7921): 217–225) to generate a stable "output signal" on a plasmid.

We thank the reviewer for this insightful suggestion. We agree that integrating Detectron with molecular recording systems such as Retro-Cascorder would be a powerful extension, enabling permanent genomic recording of Detectron-derived DNA barcodes. Because Detectron converts RNA inputs into sequence-encoded DNA, it is in principle well suited for coupling to CRISPR–Cas–based acquisition systems (e.g., Cas1–Cas2) or other integration-based recording frameworks to capture transcription events in a durable genomic format.

In the present study, we focused on establishing Detectron as a programmable RNA biosensor and defining the design principles governing its performance. We view the integration of Detectron with molecular recording systems as an important future direction that would enable long-term and cumulative recording of RNA-dependent processes, and have added discussion of this possibility in the revised manuscript:

Another potential use of this DNA could come in molecular recorders, which have relied on promoter-driven expression to generate the input to the ‘write’ operation, using tools such as CRISPR–Cas-based genome editing^{31–34}, base editors^{35,36}, integrases^{37–39}, or site-specific recombinases^{40,41}—to encode biological events indirectly. Detectrons bypass this limitation by enabling programmable RNA sensing and immediate DNA barcode synthesis upon target transcript detection. This direct RNA-to-DNA conversion introduces a new modality for RNA-responsive devices: transient RNA states can now be captured as durable, sequence-encoded DNA records, readily retrievable by sequencing. Such DNA outputs could, in principle, be coupled to CRISPR–Cas acquisition systems or other integration-based recording platforms to enable permanent genomic encoding of RNA-dependent events.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have done a reasonably thorough job responding to the reviewer 1 and 2 comments and have improved the manuscript. In particular, the additional results showing Detectron operation with three other retrons demonstrate the generality of the approach and the multiplexing capacity has been expanded to six output barcodes with a ratiometric detection approach, also strengthening the work. The revisions have also more clearly articulated the work’s conceptual advance and its positioning compared to existing technologies. I recommend publication if the issues below are addressed:

1. Extended Data Fig. 12 provides the fold change in Detectron response across MOI levels, demonstrating that MOI > 1 is required for a detectable signal. The authors should discuss the limitations of the MOI > 1 condition for practical use of the Detectrons. When varying the MOI, was the number of bacteria kept constant and the number of phage varied or vice versa?
2. The revised manuscript is still lacking in the information necessary to replicate its results. The Supplementary Tables only seem to provide the sequences for the Detectrons in the high-throughput library. The sequences for all the other Detectrons demonstrated in the paper, and their corresponding trigger sequences, should be provided. Will important plasmids be made available on Addgene? The manuscript also does not provide enough information to construct the plasmids used. What RBS was used? Was transcription of the left and right cassettes (e.g. pJH.17) carried out in the same direction or bidirectionally? Plasmid maps for the major types of plasmids used should be provided.
3. For the ratiometric detection experiments, the authors should describe why Ebu1 was selected over the other retron options.
4. There is no information provided on how the Detectron designs were changed, if at all, when moving them to the Cxx1, Eas1, and Ebu1 retron systems. The authors should provide descriptions of how this was done. Diagrams similar to Extended Data Fig. 1 would also be important, particularly if key changes to the Detectron design were required.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The authors have done a good work on the revision, and have addressed well my previous comments. I only have a few minor comments below that would further strengthen the manuscript.

Minor comments :

1. The authors should provide complete sequence details or an Addgene deposit ID for key plasmid constructions/genetic parts, rather than relying solely on construct annotations.
2. Page 3, “The high selectivity for specific bacterial species allows phage...”, ‘phage’ should be revised to phages.
3. Page 4, “To address this limitation, we developed Detectrons, a multiplexable biosensor that detects RNA and generates DNA as output signal”, ‘output signal’ should be revised to an output signal.
4. Page 5, “To test our design, we utilized a previously developed Zika virus RNA-detecting toehold switch¹¹ and retron-

- Eco1 (formerly known as ec86), the first characterized retrons in *Escherichia coli* (*E. coli*)", 'retrons' should be retron.
5. Page 8, "Complete variant and barcode sequences are provided in Supplementary Tables S4"; 'Tables' should be revised to Table. Please ensure this is the correct citation, since I did not find Table S4 in the Supplementary files.
6. Page 8, "Three separate transformations of each library serve as biological replicates", 'serve' should be revised to served.
7. Page 11, "yielded a statistically significant improvements in model performance", 'a' should be removed.

(Remarks on code availability)

Author Rebuttal letter:

We thank the reviewers for their positive assessment and for the additional minor comments on our revised manuscript. We have addressed each point below and made the corresponding changes to the text, figures, and supplementary materials. We believe these revisions have further improved the clarity and completeness of the manuscript. A point-by-point response to each comment follows.

Reviewer #1 (Remarks to the Author):

The authors have done a reasonably thorough job responding to the reviewer 1 and 2 comments and have improved the manuscript. In particular, the additional results showing Detectron operation with three other retrons demonstrate the generality of the approach and the multiplexing capacity has been expanded to six output barcodes with a ratiometric detection approach, also strengthening the work. The revisions have also more clearly articulated the work's conceptual advance and its positioning compared to existing technologies. I recommend publication if the issues below are addressed:

1. Extended Data Fig. 12 provides the fold change in Detectron response across MOI levels, demonstrating that MOI > 1 is required for a detectable signal. The authors should discuss the limitations of the MOI > 1 condition for practical use of the Detectrons. When varying the MOI, was the number of bacteria kept constant and the number of phage varied or vice versa?

We thank the reviewer for this important point. In the MOI titration experiment, the number of bacteria was kept constant and the amount of T7 phage was varied to achieve the indicated MOIs. We have clarified this in the revised Methods and figure legend.

We agree that the requirement for MOI > 1 represents an important limitation of the current Detectron implementation. Under the conditions tested, barcode enrichment was not clearly detectable at MOIs of 0.5 or 1, but became apparent at MOIs above 1 and increased further at higher MOIs. Thus, the current T7-targeting Detectron is best suited for controlled phage-challenge experiments in which infection levels are sufficiently high to produce detectable target-transcript-dependent barcode output. This includes comparative assays such as pooled host-susceptibility profiling across bacterial genotypes, rather than applications requiring detection of rare infection events or very low phage burden.

We have revised the manuscript to clarify that Extended Data Fig. 12, now Extended Data Fig. 8 in the revised manuscript, defines the MOI-dependent response under the tested assay conditions, rather than establishing a general limit of detection:

To further define the response of Detectron across infection levels, we performed an additional T7 phage infection experiment across a range of MOIs (0.5, 1, 2, 5, and 10), keeping the initial bacterial density constant while varying the amount of input phage. Barcode enrichment increased with increasing MOI, demonstrating an MOI-dependent Detectron response (Extended Data Fig. 8). Under the conditions tested, barcode enrichment was not observed at MOIs of 0.5 and 1 relative to the no-infection control, but became detectable at higher MOIs and increased further at an MOI of 10. These results define the MOI-dependent response of the T7 Detectron under the tested assay conditions and indicate that the current implementation is best suited for controlled phage-challenge experiments with sufficient infection burden, rather than detection of rare infection events or very low phage abundance.

Future improvements in target-site selection, switch architecture, retron expression, barcode recovery, and sequencing depth may improve sensitivity at lower infection levels.

2. The revised manuscript is still lacking in the information necessary to replicate its results. The Supplementary Tables only seem to provide the sequences for the Detectrons in the high-throughput library. The sequences for all the other Detectrons demonstrated in the paper, and their corresponding trigger sequences, should be provided. Will important plasmids be made available on Addgene? The manuscript also does not provide enough information to construct the plasmids used. What RBS was used? Was transcription of the left and right cassettes (e.g. pJH.17) carried out in the same direction or bidirectionally? Plasmid maps for the major types of plasmids used should be provided.

We thank the reviewer for these comments. We have substantially expanded the sequence and plasmid information provided.

We address each point below.

a. Detectron and trigger sequences. We agree that all Detectrons demonstrated in the paper, not only those in the

high-throughput library, should be fully documented. We have added a new Supplementary Table 4 providing the complete sequences of every Detectron used in the study, together with their corresponding trigger sequences.

b. Availability on Addgene. Thirty key plasmids used in this study, spanning all major plasmid types demonstrated, will be made available on Addgene. Addgene plasmid IDs have been added to Supplementary Table 3. Full plasmid maps and complete plasmid sequences will be available through the corresponding Addgene deposit pages.

c. RBS. The RBS used in our constructs is the T7 gene 10-derived ribosome binding site (Shine-Dalgarno core AAGGAG).

d. Cassette orientation. For all plasmids containing left and right cassettes (e.g., pJH.017), transcription of the two cassettes was carried out bidirectionally (divergently). We have clarified this in the text and indicated the orientation on the plasmid maps.

3. For the ratiometric detection experiments, the authors should describe why Ebu1 was selected over the other retron options.

We thank the reviewer for the question. Ebu1 was selected as a representative retron for the ratiometric detection experiments; it was among the constructs we established first in our pipeline and performed reliably, so we used it for these proof-of-concept experiments. We did not select it based on superior performance relative to the other retrons, and we would expect the ratiometric approach to be generalizable across the retron options characterized in this study.

4. There is no information provided on how the Detectron designs were changed, if at all, when moving them to the Cxx1, Eas1, and Ebu1 retron systems. The authors should provide descriptions of how this was done. Diagrams similar to Extended

Data Fig. 1 would also be important, particularly if key changes to the Detectron design were required.

We thank the reviewer for this suggestion. The Detectron-V2 toehold-switch module (sequence and architecture) was unchanged across the Cxx1, Eas1, and Ebu1 systems; only its integration into each retron ncRNA differed, and no changes to the switch design itself were required.

Although the Cxx1, Eas1, and Ebu1 ncRNAs adopt different overall structures, they share the conserved a1 and a2 elements

that form a duplex. To generate each Detectron, we removed the native a1 and a2 from the wild-type retron ncRNA and inserted the Detectron-V2 toehold-switch sequence immediately 5' of the priming G. The reverse complement of the a2 sequence was then used as the new a1, preserving the a1/a2 duplex required to prime reverse transcription. This same strategy was applied identically to all three retrons.

We have added a new figure (Supplementary Fig. 5) illustrating this design, showing the wild-type retron ncRNA, the Detectron-V2 toehold switch, and the resulting integrated Detectron for the Cxx1, Eas1, and Ebu1 systems.

Reviewer #3 (Remarks to the Author):

The authors have done a good work on the revision, and have addressed well my previous comments. I only have a few minor comments below that would further strengthen the manuscript.

Minor comments :

1. The authors should provide complete sequence details or an Addgene deposit ID for key plasmid constructions/genetic parts, rather than relying solely on construct annotations.

We agree. We have deposited the key plasmids with Addgene, and the corresponding Addgene IDs — together with detailed construct annotations — are now listed in Supplementary Table 3. Full plasmid maps and complete plasmid sequences will be available through the corresponding Addgene deposit pages. Complete sequences of all Detectrons and trigger RNAs used in this study are provided in Supplementary Table 4. Both are referenced in the “Plasmid construction and architecture” Methods section.

2. Page 3, “The high selectivity for specific bacterial species allows phage...”, ‘phage’ should be revised to phages.

Corrected: “Their high selectivity for specific bacterial species allows phages...”

3. Page 4, “To address this limitation, we developed Detectrons, a multiplexable biosensor that detects RNA and generates DNA as output signal”, ‘output signal’ should be revised to an output signal.

We thank the reviewer for pointing this out. This sentence has been removed from the revised manuscript, and the phrase

“output signal” no longer appears in this context.

4. Page 5, “To test our design, we utilized a previously developed Zika virus RNA-detecting toehold switch¹¹ and retron-Eco1 (formerly known as ec86), the first characterized retrons in *Escherichia coli* (*E. coli*)”, ‘retrons’ should be retron.

Corrected: “...and retron-Eco1 (formerly known as ec86), the first characterized retron in *Escherichia coli* (*E. coli*).”

5. Page 8, “Complete variant and barcode sequences are provided in Supplementary Tables S4”; ‘Tables’ should be revised to Table. Please ensure this is the correct citation, since I did not find Table S4 in the Supplementary files.

We thank the reviewer for catching this. We have corrected “Tables” to the singular and updated the citation. The complete variant and barcode sequences are now provided in Supplementary Table 1. Because this table contains 23,520 variant-barcoded sequences and is too large to embed in the Supplementary Information PDF, it is provided in the separate Supplementary Table Excel file. We have confirmed that the reference is now present and correctly numbered. We apologize for the omission in the previous version.

6. Page 8, “Three separate transformations of each library serve as biological replicates”, ‘serve’ should be revised to served.

Corrected: “Three separate transformations of each library served as biological replicates”

7. Page 11, “yielded a statistically significant improvements in model performance”, ‘a’ should be removed.

Corrected: “...yielded statistically significant improvements in model performance.”

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>