End-to-end computational approach to the design of RNA biosensors for detecting miRNA biomarkers of cervical cancer

Priyannth Ramasami S. Baabu [b,c,1], Shivaramakrishna Srinivasan [a,d,1], Swetha Nagarajan [b], Sangeetha Muthamilselvan [a], Thamarai Selvi [a], Raghavv R. Suresh [b], Ashok Palaniappan [a,*]

[a] *Department of Bioinformatics, School of Chemical and Biotechnology, SASTRA Deemed University, Thanjavur, Tamil Nadu 613401, India*
[b] *Department of Bioengineering, School of Chemical and Biotechnology, SASTRA Deemed University, Thanjavur, Tamil Nadu, 613 401, India*
[c] *School of Advanced Materials Science and Engineering, Sungkyunkwan University (SKKU), Suwon, 16419, Republic of Korea*
[d] *Life Sciences Division, TCS Research, Tata Consultancy Services, Chennai, Tamil Nadu, 600042, India*

ARTICLE INFO

ABSTRACT

Cervical cancer is a global public health subject as it affects women in the reproductive ages, and accounts for the second largest burden among cancer patients worldwide with an unforgiving 50% mortality rate. Relatively scant awareness and limited access to effective diagnosis have led to this enormous disease burden, calling for point-of-care, minimally invasive diagnosis methods. Here, an end-to-end quantitative unified pipeline for diagnosis has been developed, beginning with identification of optimal biomarkers, concurrent design of toehold switch sensors, and finally simulation of the designed diagnostic circuits to assess performance. Using miRNA expression data in the public domain, we identified miR-21–5p and miR-20a-5p as blood-based miRNA biomarkers specific to early-stage cervical cancer employing a multi-tier algorithmic screening. Synthetic riboregulators called toehold switches specific to the biomarker panel were then designed. To predict the dynamic range of toehold switches for use in genetic circuits as biosensors, we used a generic grammar of these switches, and built a neural network model of dynamic range using thermodynamic features derived from mRNA secondary structure and interaction. Second-generation toehold switches were used to overcome the design challenges associated with miRNA biomarkers. The resultant model yielded an adj. R^2 ~0.71, outperforming earlier models of toehold-switch dynamic range. Reaction kinetics modelling was performed to predict the sensitivity of the second-generation toehold switches to the miRNA biomarkers. Simulations showed a linear response between 10 nM and 100 nM before saturation. Our study demonstrates an end-to-end computational workflow for the efficient design of genetic circuits geared towards the effective detection of unique genomic/nucleic-acid signatures. The approach has the potential to replace iterative experimental trial and error, and focus time, money, and efforts. All software including the toehold grammar parser, neural network model and reaction kinetics simulation are available as open-source software (https://github.com/SASTRA-iGEM2019) under GNU GPLv3 licence.

1. Introduction

Cervical cancer is the second most common cancer affecting women worldwide [1–3], with 20% of cases in India [4,5]. Most cervical cancers could be attributed to the HPV16 and HPV18 strains of Human Papilloma Virus (HPV) [1,2]. Lifestyle factors also contribute to the etiology of the disease [2]. Cervical cancer tumorigenesis involves infection of the metaplastic epithelium by HPV at the cervical transitional zone,

followed by viral persistence and the progression to pre-cancerous epithelial cells, thus nucleating the initiation of cancer [6]. Pap smear test is the gold-standard diagnostic method, but it is invasive, expensive, and time-consuming [4,7]. Alternative testing strategies are necessary [8–10], and micro-RNAs (miRNAs) circulating in the blood have emerged as effective biomarkers in many cases [11–14]. Using data-driven methods for the identification of reliable early-stage diagnostic/prognostic biomarkers, coupled with the concurrent design of

Table 1

Iterative Design, Build, Test & Learn (DBTL) cycles deployed for developing toehold-switch biosensors specific for a panel of miRNA biomarkers of early-stage cervical cancer.

DBTL phase	Iteration 0: Biomarker signature development	Iteration 1: Optimal toehold switch design	Iteration 2: Reaction network modelling of circuit performance
DESIGN	TCGA data-driven differential expression analysis, survival analysis	Machine learning to design optimal toeholds for identified biomarkers	Modelling the designed circuit using chemical and biochemical kinetics
BUILD	Cancer vs. control linear models, univariate K-M plots, multivariate Cox regression models (using R)	Feature engineering of toehold datasets using ViennaRNA and in-house Python and shell scripts	Generalising the model to second-generation toeholds with anti-miRs, to inform the design of experiments
TEST	Models evaluated using p-values and log fold-change of differential expression, and p-values of prognosis models.	Multivariate regression model of toehold efficacy evaluated using R^2 goodness of fit (using R)	Validation of predictive modelling
LEARN	Filtering to four biomarkers, then to three, and finally to two	Optimized toehold switch sequences	Sensitivity, onset of saturation, and specificity of each sensor

biosensors offers potential for the development of a new generation of molecular diagnostics in a single workflow [15,16].

Synthetic biology is delivering on its promise of harnessing nature's inherent diversity for the betterment of human health [17]. Bio-electronic diagnostic devices comprise an integrated single-unit reaction chamber housing the sensor element with the necessary reagents for analyte detection, and the corresponding transduction module for providing a reaction readout with optical, piezoelectric, or electrochemical means [18–20]. Toehold switch riboregulators are synthetic mRNA elements, which function by occluding the ribosome from translating a downstream gene [21]. Upon base pairing with the trigger RNA sequence, the toehold structure of the switch unfolds, permitting expression of a reporter gene. Such sensors could sense virtually any RNA sequence with excellent biosensor properties, namely specificity, modularity, and orthogonality [21], and could be freeze-dried on a microfluidic platform for biomarker detection using a cell-free colorimetric assay [22]. The toehold-switch concept was extensively utilised

in the design of a biosensor for detecting Zika virus infection [23,24]. Takahashi et al. designed a toehold switch sensor that could detect the conserved region of *C. difficile* in a paper based platform [25]. In this work, we have formulated an alternative testing strategy for cervical cancer based on data-driven identification of biomarker miRNAs, design of cognate toehold-based sensors, followed by *in silico* modelling of sensor-circuit performance. The main contributions are twofold: (1) an end-to-end unified computational pathway from disease omics to biosensor design; and (2) application of this pathway to yield potential diagnostics/prognostics for cervical cancer. Supplementary Information accompanying the manuscript could be found at https://doi.org/10.6084/m9.figshare.14915619.v5, and the software required to reproduce the workflow and use the tools described herein is available open-source under GNU GPLv3 at https://github.com/SASTRA-iGEM2019.

2. Methods

The overall strategy is outlined in Table 1 and discussed in detail below.

2.1. Biomarker panel development

The Cancer Genome Atlas (TCGA) provided the miRNA-Seq dataset with 309 cervical adenocarcinoma samples and 3 tumor-normal matched controls [26]. The stage information was extracted from the associated clinical data using the attribute 'patient.stage_event.clinical_stage'. There were 163 stage-I, 70 stage-II, 46 stage-III and 21 stage-IV samples. The patient sample barcode encoded in variable *hybridisation_REF* was parsed to annotate the sample as cancer or normal.

RSEM-normalized Illumina HiSeq miRNASeq gene expression data [27] was $\log_2$ transformed, and analyzed for differential expression using R (https://www.r-project.org/) *limma* package [28]. Linear modelling of stage-annotated gene expression matrix was performed, followed by empirical Bayes adjustment to obtain moderated t-statistics. To account for multiple hypothesis testing and the false discovery rate, the p values of the F-statistic of linear fit were adjusted using the BH method [29]. Differentially expressed miRNAs were considered significant if the absolute log-fold changes >1.5x and the p-value was <0.05 [30]. A tiered contrast analysis for identifying stage-specific DE miRNAs was then carried out (see the Methods section in Ref. [30]). Finally, the association between the significant early-stage DE miRNAs and the overall survival of a patient was evaluated by univariate Cox proportional hazards regression analysis, and the significant prognostic miRNA
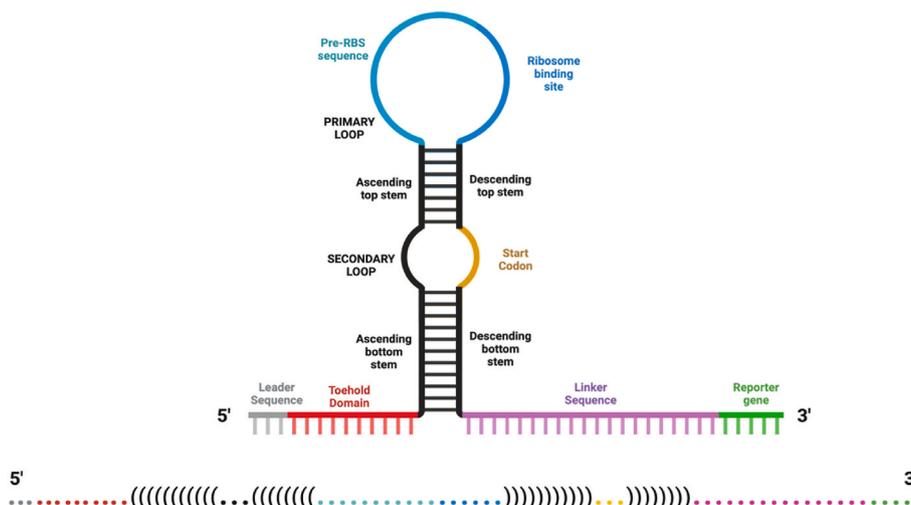
Fig. 1. Anatomy of a toehold switch, showing its main features; and below, translating the secondary structure into a generic dot-bracket representation of the toehold switch grammar. The following criteria were used in the rational design of toehold sequences.
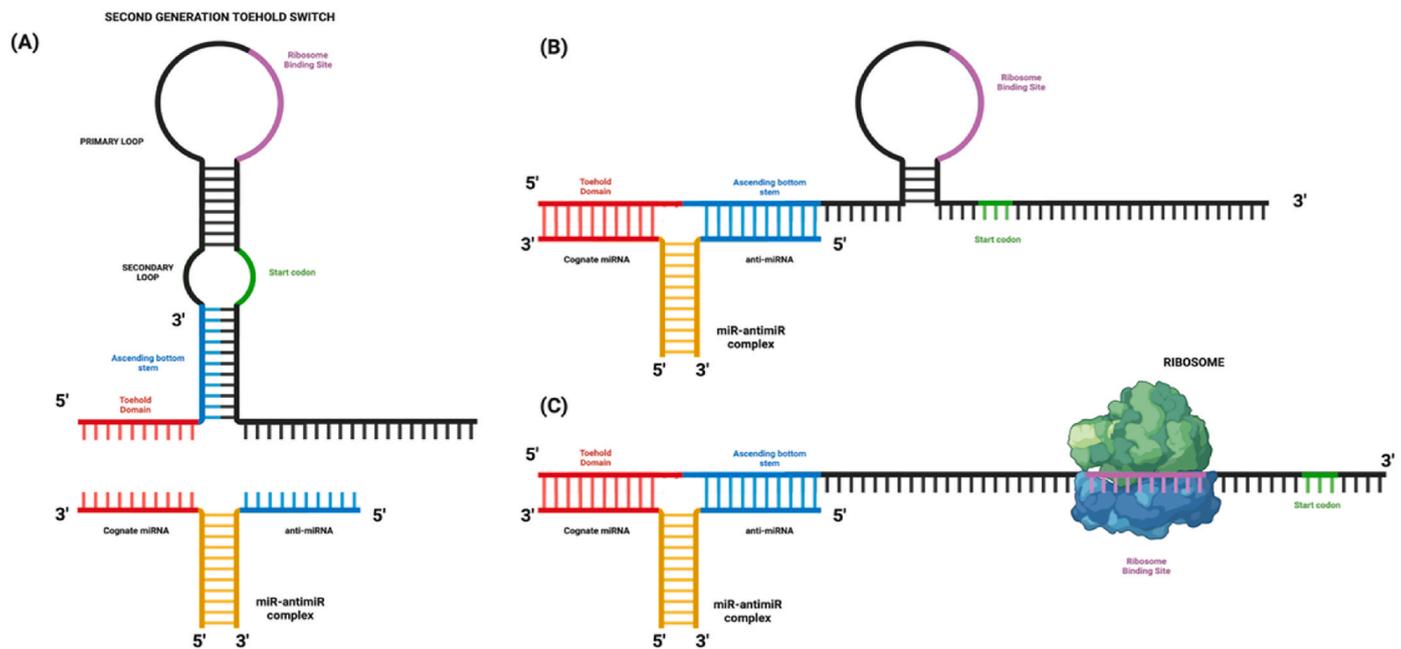
Fig. 2. Second-generation toehold switches. (A) Illustration of a second generation toehold switch construct with the hybrid trigger complex between the miRNA and antimiR. (B) Binding of the hybrid trigger with the toehold domain, unspooling the bottom stem, and (C) subsequent linearization of the toehold switch releasing the RBS from occlusion, and yielding translation of the reporter gene. Created with BioRender.com.

biomarkers were identified. Survival analysis was done using the R survival package [31].

2.2. Toehold switch design

2.2.1. Toehold switch grammar

Toehold switches act as sensors for exact sequences of RNA and could be designed to hybridize with cognate molecules. The toehold structure is marked by a hairpin loop sequestering the RBS, an internal loop containing the start codon, and two base paired stems, and functions as a regulatory-switch of translation by occluding the RBS in the absence of the cognate trigger molecule ('ground state') [21] (Fig. 1). Upon binding of the cognate miRNA (i.e, the trigger sequence) to the linear toehold domain, the bottom stem unravels and the double-looped structure of the toehold switch collapses to a linear structure, exposing the sequestered start codon to expression of the reporter gene ('active state'). Each toehold switch is a modular entity with distinct sequence domains engineered *de novo* for a particular trigger RNA. The dot-bracket notation encodes a representative grammar of RNA secondary structure, '.' representing unpaired bases, and matched-parentheses '(' and ')' representing paired bases (Fig. 1, bottom). Regular expressions could be used to parse dot-bracket notation. An in-house Python script based on regular expressions and ViennaRNA utilities [32] was developed to parse the grammar of toehold switches into their constituent domains (https://github.com/SASTRA-iGEM2019), accounting for the following scenarios:

1. Wobble base-pair between G-U in the secondary loop containing the start codon.
2. Base-paired toehold domains
3. Base-paired linker regions
4. Possible unpaired bases in the top stem
1. The sequence of the toehold domain depends on the trigger complex and was at least 10 nts to allow effective binding of the trigger to the switch.
2. The descending bottom stem of the toehold switch was about 12 nucleotides, to maintain switch stability in the OFF state and ORF integrity.

3. The canonical RBS sequence AGAGGA was used.
4. The primary loop was made adenine-rich to maintain a larger loop structure.

The toehold-switch design and grammar are anchored with the following parts: RBS, 21-nt linker, start codon region, and Shine Dalgarno sequence. The variable domain is the trigger binding region (inclusive of the toehold domain and the ascending bottom stem) which obeys base-complementarity with the trigger.

In this case, the triggers of interest are mature miRNAs about 22 nts long, which might include a stop-codon trinucleotide towards their 5′ sequence. The presence of stop-codon subsequences could halt the expression of the downstream GFP, and pose a problem to the use of toehold-switches. Second-generation toehold switches are one solution to this problem (Fig. 2(A)) [33]. In addition to binding the miRNA, the toehold switch is required to bind another molecule, the anti-miRNA (antimiR) that is complementary to 12 nucleotides at the 3′-end of the toehold switch (Fig. 2(B, C)). Essentially, the trigger is represented by a hybrid molecule with a base-paired region and two free ends. The design of the antimiR sequence could be used in improving the trigger binding event.

2.2.2. Machine learning of toehold switch efficacy

The sensitivity of the toehold switch is described by its dynamic range, defined as the ratio between the maximum and minimum measurable intensities of reporter protein (typically GFP) expression, also known as ON/OFF ratio [21]. The dynamic range of a toehold designed for exact sequences of trigger RNA is the key indicator of its effectiveness, and is generally unknown for new toehold designs. The problem can be addressed using supervised machine learning on toehold-sequence datasets with available dynamic range responses. A previous model developed by CUHK iGEM 2017 offered a web tool for predicting the dynamic range of toehold switches with modest goodness of fit (unadjusted R^2 ~0.04–0.16) [34]. To address this lacuna, we constructed multiple models that utilize the generic grammar of the toehold switch discussed above. A dataset of the toehold switches and their dynamic range responses was formed by combining the studies of Green et al. [21] and Pardee et al. [22,23]. A total of 228 toehold switch
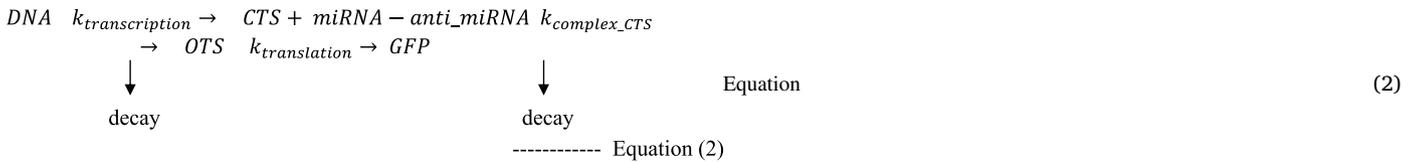
sequences with ON/OFF ratios were obtained, consisting of:

(1) first-generation switches with moderate dynamic range (168 instances) [21].
(2) forward-engineered switches with significantly higher dynamic range ratios (13 instances) [21].
(3) toehold switches engineered for detecting Zika virus (47 instances) [23].

The toehold sequence of each instance was parsed into its subdomains using our software. Sequence features were considered but rejected in favor of the more informative structure-based features [34], which were computed using ViennaRNA utilities. These features included the overall switch minimum free energy (MFE), RBS-linker MFE, bottom-region MFE, and the net MFE of the second-generation

$$miRNA + anti\_miRNA \rightleftharpoons miRNA - anti\_miRNA \ \text{(upstream reaction)}$$

------------ Equation (1)
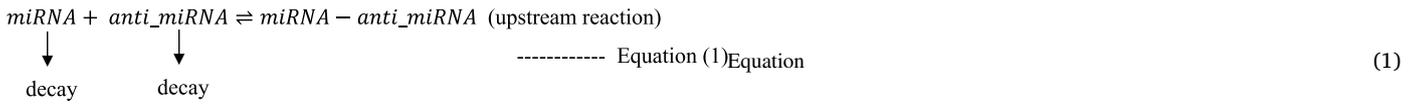
switch-trigger complex (which was calculated as the difference in MFE between the final and initial states). The overall switch MFE determines the switch stability in the OFF state, the RBS-linker MFE is a thermodynamic proxy of the translation initiation rate, and the net MFE mea-

$$DNA \ k_{transcription} \rightarrow \ CTS + miRNA - anti\_miRNA \ k_{complex\_CTS}$$
$$\rightarrow \ OTS \ k_{translation} \rightarrow GFP$$

decay   decay

------------ Equation (2)

$$GFP \ \overset{k_{maturation}}{\rightarrow} \ Mature \ GFP (downstream \ reaction)$$

sures the energetic cost of the binding events taking place during toehold-mediated strand displacement.

In this manner, a dataset of engineered features and the corresponding toehold efficacies was compiled. Following an 80:20 train:test

$$\frac{d[miRNA]}{dt} = -k_{miR\_antimiR\_f}[anti\_miRNA][miRNA] + k_{miR\_antimiR\_b}[miRNA - anti\_miRNA] - k_{decay}[miRNA]$$

$$\frac{d[anti\_miRNA]}{dt} = -k_{miR\_antimiR\_f}[anti\_miRNA][miRNA] + k_{miR\_antimiR\_b}[miRNA - anti\_miRNA] - k_{decay}[anti\_miRNA]$$

$$\frac{d[miRNA - anti\_miRNA]}{dt} = -k_{miR\_antimiR\_f}[anti\_miRNA][miRNA] + k_{miR\_antimiR\_b}[miRNA - anti\_miRNA]$$

split, the features were normalized. Two regressor techniques were used to learn the relationship of the toehold-switch dynamic range with the feature space. One, a multivariate linear regression model was built with the selected features on the train data, and evaluated on the test data for significance and goodness-of-fit, using Python scikit-learn (https://scikit-learn.org/). Second, a neural network with one hidden layer

of 6 neurons, with L1 regularization and early stopping to control overfitting, was constructed based on the Keras framework [35].

2.3. Reaction network modelling

The chemical and biochemical kinetics model was developed by adding the following features to the first generation toehold switch model developed by CLSB-UK iGEM 2017 [36]:

(i) interaction of transcribed miR and ntimir oligos to form the miR-antimiR complex with two free ends, which then binds with the toehold switch
(ii) maturation kinetics of the expressed GFP.

This generalized model can be presented as:

$$\text{Equation} \tag{1}$$

Here, $k_{miR\_antimiR\_f}$ and $k_{miR\_antimiR\_b}$ are the rate constants for forward and backward reactions, respectively.

$$\text{Equation} \tag{2}$$

$$\tag{3}$$

The corresponding ODEs can be given by:

$$\frac{d[miRNA - anti\_miRNA]}{dt} = -k_{complex\_CTS}[CTS][miRNA - anti\_miRNA]$$

$$\frac{d[CTS]}{dt} = k_{transcription}[DNA] - k_{complex.CTS}[CTS][miRNA - anti.miRNA] - k_{decay}[CTS]$$

$$\frac{d[OTS]}{dt} = k_{complex\_CTS}[CTS][miRNA - anti\_miRNA] - k_{decay}[OTS]$$

$$\frac{d[GFP]}{dt} = k_{translation}[OTS] - k_{maturation}[GFP]$$

$$\frac{d[mature\,GFP]}{dt} = k_{maturation}[GFP]$$

The kinetic parameters for the transcription, translation and maturation reactions in the genetic circuit were identified from literature [36, 37] and provided in Supplementary File S6. The kinetics for the miR-antimiR complex formation were modelled assuming a steady state with negligible backward flux, so that the equilibrium constant approximates the forward rate constant at equilibrium. From equation (1), the equilibrium constant for the complex formation is defined as:

$$k_{eq} = \frac{[hsa\_miR\_21\_5p - anti\_hsa\_miR\_21\_5p]}{[hsa\_miR\_21\_5p][anti\_hsa\_miR\_21\_5p]} \quad (4)$$

NUPACK Analysis tool was used to obtain the equilibrium concentration of the miR, ntimir and their complex [38]. The chemical and biochemical kinetics were simulated using MATLAB R2017b (https://www.mathworks.com/).

3. Results

3.1. Omics analysis

The top 100 miRNAs of the linear model were screened for log-fold-change with respect to the control samples, stage-specificity, and significance of the trend, and finally corroborated with the literature evidence. This analysis yielded four DE miRNAs (Table 2) [12,39]. It could be seen from the heat map and violin plots that hsa-miR-21–5p is the most significantly differentially expressed miRNA (Fig. 3). The detailed results of the differential expression analysis and linear modelling are provided in Supplementary Files S1 and S2, respectively. It is interesting to note that all four miRNAs are upregulated across all stages of cervical cancer, suggesting possible role as oncomiRs and use in molecular diagnostics (Fig, S4, Supplementary File S3).

3.2. Survival analysis

The differential expression profiling of miRNAs associated with early-stage cervical cancer identified potential biomarkers, which were further investigated using survival analysis with the Cox proportional hazards model and Kaplan-Meier curves. The biomarkers were subjected to univariate survival analysis and the model significance was estimated using the log-rank test [32]. Fig. 4 shows the K-M plots of the univariate survival analysis for the four differentially expressed miRNAs. It could be seen that hsa-miR-29a-3p was insignificant (p-value > 0.05), and was consequently dropped from the hazard ratio estimation (Table 3).

Next, we constructed a survival risk-score using multivariate Cox regression model [19], to obtain an equation for the survival risk score (SRS).

and low-risk groups using an optimal cut point identified with the maxstat statistic of *survminer* R package. Overall survival curves were generated using the K-M method, and two-sided log-rank tests were used to compare the differences in overall survival between the risk groups. The K-M analysis shown in Fig. 5 (a), indicated that the biomarker panel with three miRNAs was significant (p-value < 0.001). Considering the significance attributed to hsa-miR-21–5p in the literature [12,39–44], we then evaluated a biomarker panel of just the two miRNAs, hsa-miR-20a-5p and hsa-miR-21–5p,. The K-M curve shown in Fig. 5 (b), indicated that the significance of this panel was also significant (p-value < 0.001). However, a multivariate model of hsa-miR-21–5p and hsa-miR-200a-5p was not significant (p-value ~ 0.1). The univariate model shown in Fig. 4(b) and the form of eqn (5) both suggest that the overexpression of miR-200a-5p might have a protective effect against cervical cancer, and its expression might be a response to the cancer progression, making it inefficacious as a marker of cervical cancer. From the above results on differential expression and prognosis, we proposed to consider the two-marker panel of hsa-miR-20a-5p and hsa-miR-21–5p as the signature of early-stage cervical cancer.

We sought to validate this identified biomarker panel with an external miRNA dataset. Towards this, we used the GEO dataset, GSE30656, of 10 normal and 37 cervical cancer patients. The dataset was subjected to stage-annotated linear modelling as in the previous case (using the protocol described in Ref. [30]), and yielded a confirmation that hsa-miR-21–5p is the top-ranked differentially expressed (upregulated) miRNA in this dataset as well (p-value < 1E-4). The details of this analysis are provided in Supplementary File S4. There is extensive literature support for hsa-miR-20a-5p as an upregulated early-stage serum biomarker of cervical cancer progression [40,41, 45–51]. To investigate the pathways activated by the identified biomarkers, we first identified their regulatory targets, and then performed an enrichment analysis of the target genes using miRabel [52] and DAVID [53], ascertaining the involvement of oncogenic pathways (Supplementary File S5). In particular, the KEGG enrichment analysis of the consensus target genes of both the biomarkers yielded significance for the MAPK signaling pathway. The two differentially expressed up-regulated miRNAs with prognostic significance constituted a reliable biomarker panel.

3.3. Toehold switch design

Two toehold switches were designed to target the biomarkers, hsa-miR-21–5p and hsa-miR-20a-5p. Table 4 shows the domain-by-domain decomposition of each of these toehold switches. The folding of the transcribed RNA of the designed sequences was predicted using ViennaRNA [32], and the predicted secondary structure was consistent with the toehold conformation (Fig. 6(a, b)). The minimum free energies of the hsa-miR-21–5p toehold switch and the hsa-miR-20a-5p toehold switch were estimated to be −21.20 kcal/mol and −19.20 kcal/mol respectively, suggesting a stable toehold conformation ready for interaction with the cognate hybrid miR-antimiR trigger.

In addition to the secondary structure and the minimum free energy, the dynamic range of the designed toehold switches is a key parameter determining the effectiveness of system. A correlation plot of all

$$SRS = (-0.5967 * expression\,value\,of\,hsa - miR - 20a - 5p) + (0.5376 * expression\,value\,of\,hsa - miR - 200a - 3p) + (-0.4736 * expression\,value\,of\,hsa - miR - 21 - 5p) \quad (5)$$

Finally, a biomarker panel of hsa-miR-20a-5p, hsa-miR-21–5p and hsa-miR-200a-5p was constructed, classifying patients into high-risk

considered features for learning the dynamic range of toehold switches is discussed in Supplementary File S6. The linear model yielded significance (p-value < 0.05) for all the six features (Supplementary File S6)

Table 2
DE miRNA biomarkers of cervical cancer. LogFC (log-fold change) and adj. p-values with respect to the control as estimated by the linear model are shown. The stage-specificity (individual log-fold changes and overall significance of contrast analysis) of the miRNA is computed using the protocol described in Ref. [30].

| DE miRNAs	logFC	Adj. P value	LFC: Stage_1 – Control	LFC: Stage_2 – Control	LFC: Stage_3 – Control	LFC: Stage_4 – Control	p.value: Stage-specificity
hsa-miR-21–5p	15.75	3E-151	2.63	2.63	2.75	2.65	4E-08
hsa-miR-20a-5p	6.60	5E-26	2.01	2.02	2.22	2.20	1E-02
hsa-miR-29a-3p	13.06	1E-70	1.99	2.31	2.18	2.64	6E-03
hsa-miR-200a-5p	3.27	9E-07	6.60	6.43	6.32	6.63	3E-17

on the training set, with an adj. $R^2 \sim 0.59$. This provided a baseline to compare the performance of the neural network regressor. The detailed architecture and hyperparameters used in the neural network could be found in Supplementary File S6. The neural network architecture was experimented with different random seeds, yielding a series of models with variable performance. Of these, we chose the top twelve models that had adj. $R^2 > 0.50$ on the test set. The best of these models achieved an adj. $R^2 \sim 0.71$. The performance of this model is shown in Fig. 7, illustrating a steady improvement in the test set metrics with the training epochs. A 10-fold cross-validation with this best-performing model yielded an adj. $R^2 \sim 0.49$. The growth of the small-sample correction to the R^2 score during k-fold cross-validation tends to quickly diminish the cross-validated goodness-of-fit. The script to build and train the neural network model is provided as a Jupyter notebook (https://github.com/SASTRA-iGEM2019/). All the top models are

deposited in the project Github repository, along with the usage notes for all the software tools developed in this work.

We then applied the top twelve neural network models to predict the efficacies of the toehold switches designed for the two identified miRNA biomarkers. For the two switches, we obtained a predicted dynamic range $\sim [88.34 \pm 16.39, 92.86 \pm 18.74]$ respectively, which indicated potential robust design of the biosensors. This part of our workflow has been bundled into a single bash script that returns the predicted efficacy for any input toehold-switch sequence, and would be useful in toehold-based synthetic biology.

3.4. Reaction modelling

We employed reaction network mass action kinetics to computationally evaluate the performance of the designed second-generation
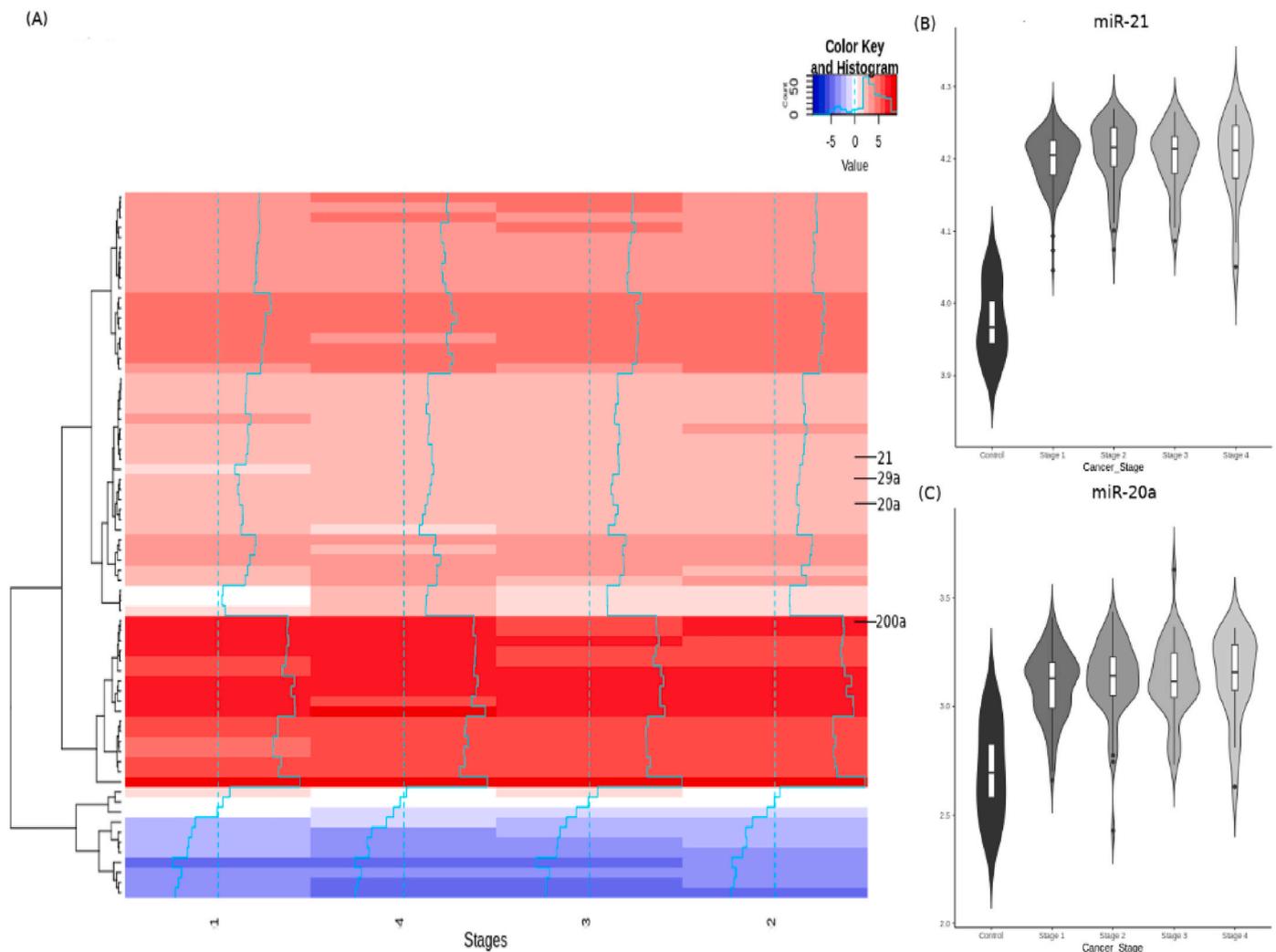
Fig. 3. Stagewise linear modelling of the top differentially expressed miRNAs from TCGA cervical cancer dataset. (A) Heat map with the four DE miRNAs considered in this study, and (B) violin plots of the two miRNAs taken forward for the final biomarker panel.
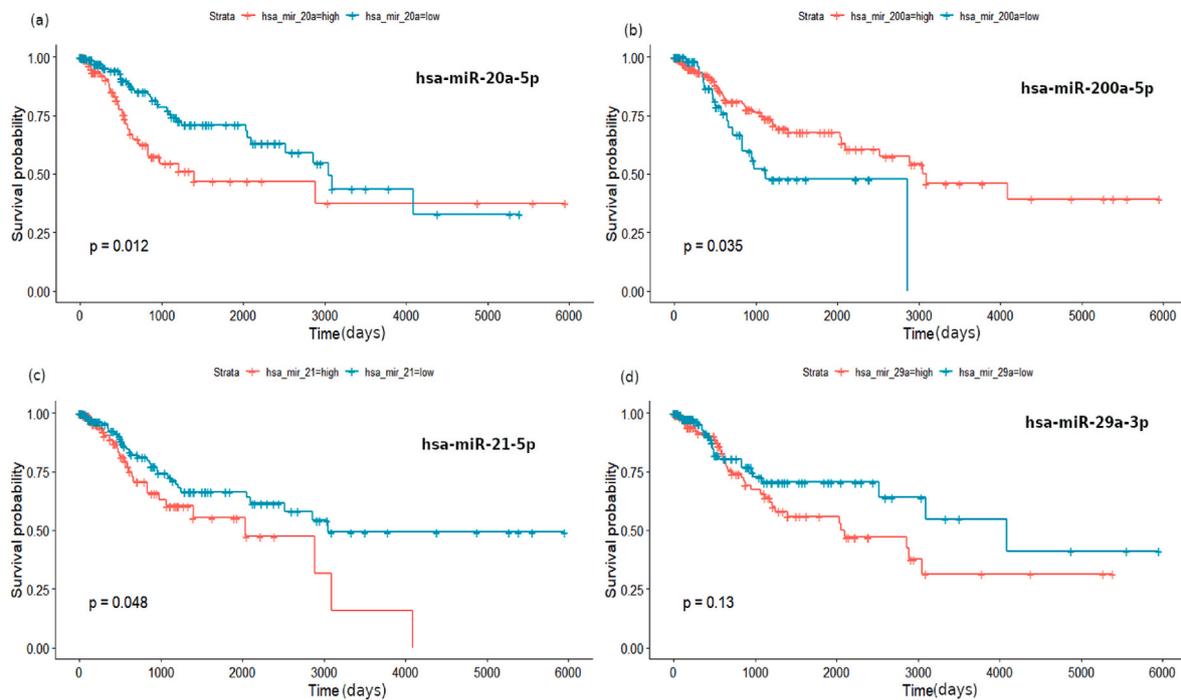
Fig. 4. K-M plot for (a) hsa-miR-20a-5p, (b) hsa-miR-200a-5p, (c) hsa-miR-21–5p and (d) hsa-miR-29a-3p. Except miR-29a-3p, the others are significant in the univariate survival model. Survival probability is adversely impacted by overexpression of miR-21–5p and miR-20a-5p, supporting their use as biomarkers and even as targets for therapy.

toehold switches as biosensors of miRNA biomarkers (available as executable SBML model in github repository, and as BioModels MODEL2107150001). As a prelude, we emulated the model proposed by iGEM CLSB-UK (2017) [36] for the first-generation toehold switches (Supplementary File S7).

3.4.1. Generalization to second generation toehold switches

The miRNA biomarkers induced a stop codon in the design of first-generation toehold switches, which could be circumvented by the use of anti-miR based second-generation toehold switches. The second-generation toehold switches were modelled by adding a reaction upstream of the miRNA-closed toehold switch binding, where the miRNA

Table 3
Hazard ratio (HR), Z-score and p-value for the chosen miRNA biomarkers. HR > 1 indicates a covariate that is positively associated with the event probability (here, negatively with survival). The Z-score is an estimate of the significance of the hazard ratio under the assumption of standard normal distribution of the log-rank test statistic.

| Gene Symbol	HR	Z-score	p-value
hsa-miR-20a-5p	0.523	−2.479	0.013
hsa-miR-200a-5p	1.808	2.076	0.038
hsa-miR-21–5p	0.595	−1.961	0.050
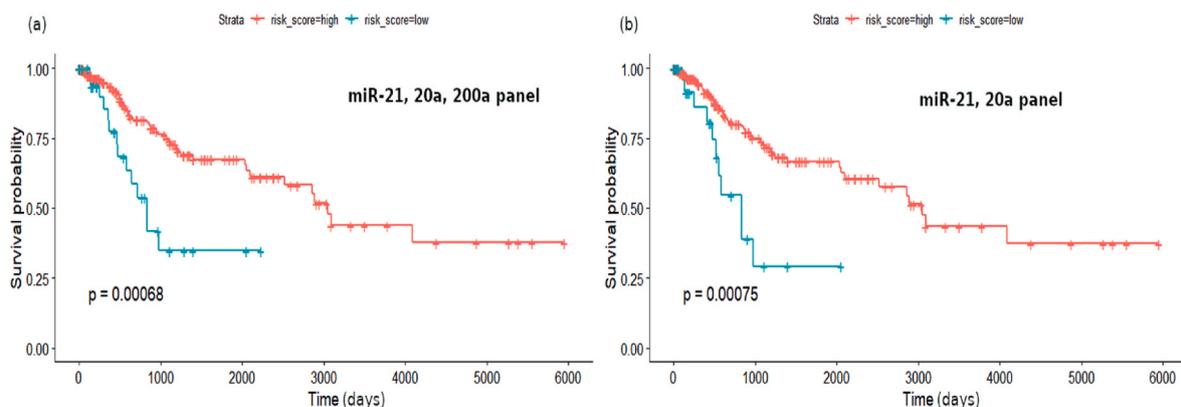
Fig. 5. Kaplan-Meier curves of multivariate models. (a) Panel with three biomarkers: hsa-miR-20a-5p, hsa-miR-21–5p and hsa-miR-200a-5p, and (b) panel with two biomarkers: hsa-miR-20a-5p and hsa-miR-21–5p.

Table 4

Domain-wise sequence design of the second generation toehold switches targeting specified miRNA biomarkers.

Domain	Biomarkers		Source
	hsa-mir-21–5p	hsa-mir-20-a-5p	
Toehold Domain	ucaacaucagc	cuaccugcacuau	Complementary to the 3′ end of the biomarker
Ascending Bottom Stem	ucguuuauc	ucguuuaguc	Complementary to the 5′ end of the anti-miR designed for the biomarker
Secondary Loop (nucleotides opposing start codon)	acg	acg	
Ascending Top Stem	uuuac	uuuac	Sequence used in Green et al. [21]
Primary Loop (RBS sequence in bold)	aaaaagaggaga	aaaaagaggaga	Chosen from the Anderson RBS family
Descending Top Stem	guaaa	guaaa	Complementary to the ascending top stem
Start Codon	aug	aug	
Descending Bottom Stem	gauaaacga	gacuaaacga	Complementary to the ascending bottom stem
Linker	aaccuggcggcagcgcaaaag	aaccuggcggcagcgcaaaag	Sequence used in Green et al. [21]
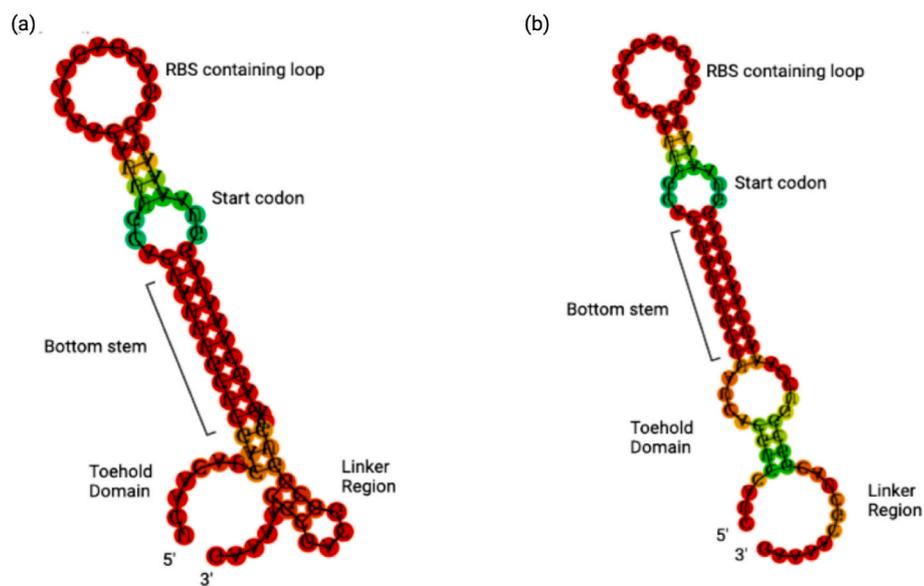
Fig. 6. ViennaRNA predicted MFE structures of the designed second-generation toehold switches. (a) hsa-miR-21–5p, (b) hsa-miR-20a-5p. It is visually clear that the toehold sequences fold into structures consistent with the canonical toehold-switch grammar with the hairpin loop, internal loop and bottom stem.
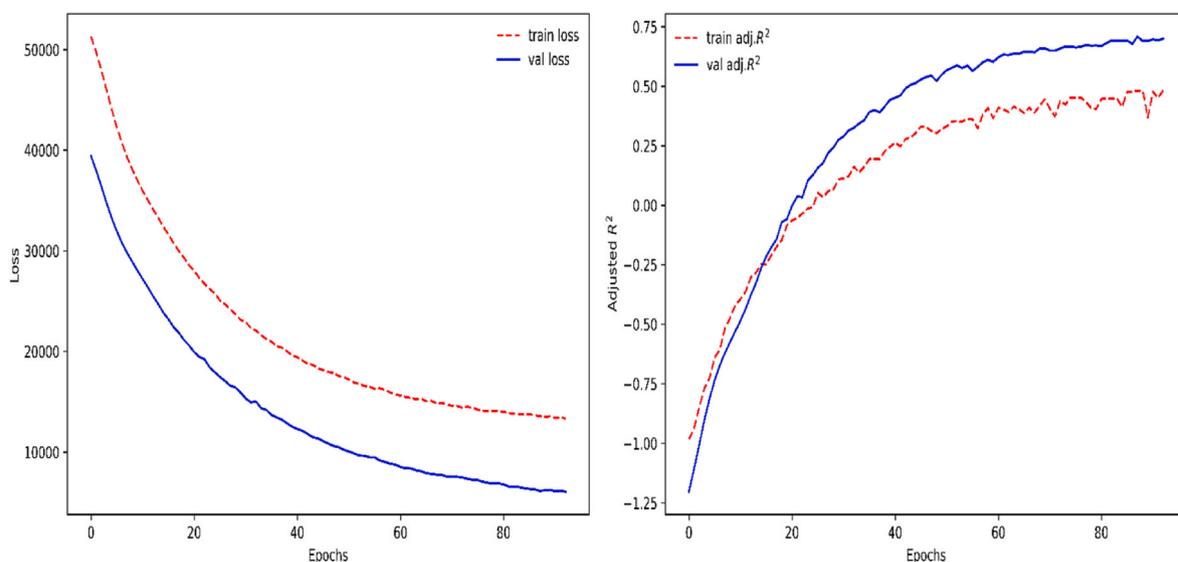
Fig. 7. Epoch tuning curves for the top-performing neural network model for the loss function and the adjusted R^2. The adj. R2 was coded as a custom metric in the model definition. The validation loss is better than the training loss, an observation that can be traced to the use of L1-regularization in the model [35]. Initial negative values for adj. R^2 indicate that the fit is worse than a 'horizontal line' in the learning manifold, but progressively trains to achieve a better fit. A steady convergence beyond epoch ~#50 is visible.
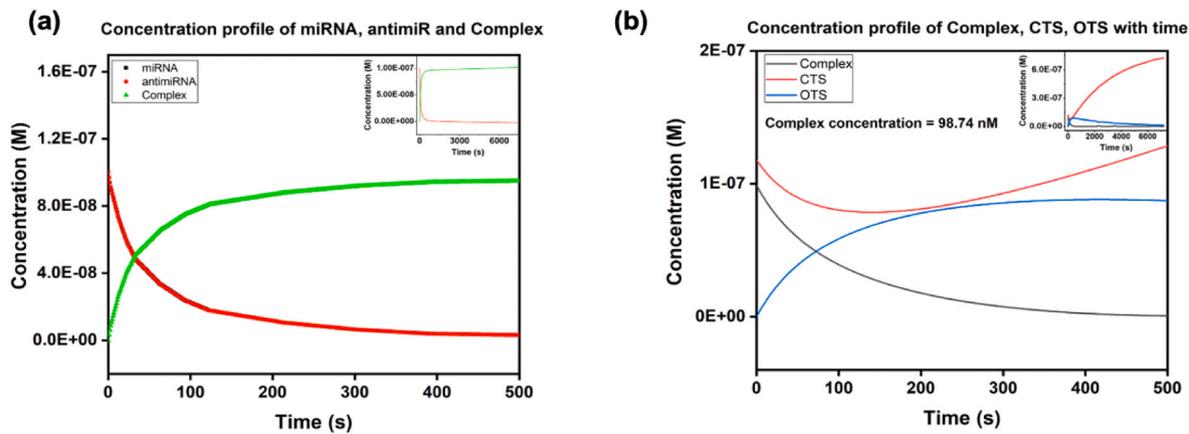
Fig. 8. Circuit kinetics modelling. (a) Concentration profiles of miR, odelli and the complex. It is seen that miR and odelli profiles are identical. (b) Concentration profiles of the complex, CTS and OTS in a 2 h time frame with a miR-antimiR binding rate constant of 10^5 nM^{-1}s^{-1}. The inset shows the profile over 7200s.
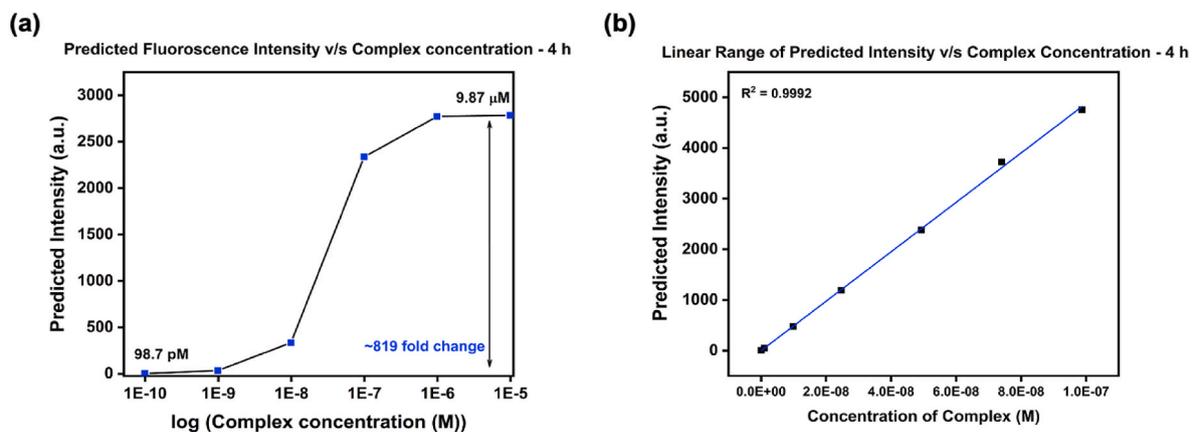

Fig. 9. Simulations show the fluorescence intensity increasing with trigger concentration. (a) An 819-fold change for trigger concentrations between 98.7 pM and 9.87 μM. (b) Detail of the linear trend for trigger concentrations between 10 nM and 100 nM.

hybridizes with the antimiR. NUPACK [38] was used to estimate the equilibrium concentrations of 100 nM hsa-miR-21–5p miRNA and 100 nM antimiR hybridizing at 18 °C to form the trigger complex (yielding 0.0264 nM, 0.0264 nM and 99.97 nM, respectively). This yielded a hybridisation rate constant $k_{miR\_antimiR\_f}$ on the order of 10^5 nM^{-1} s^{-1} (Eqn (4)). The above analysis was repeated with hsa-miR-20a-5p, and the details could be found in Supplementary File S7 – Sec. S2.

Following the binding of the complex to the toehold switch, the switch would open, enabling translation of the downstream GFP. Since GFP fluorescence is associated with a lag between production and emission, a GFP maturation reaction was considered with a maturation factor ∼0.2 min^{-1}. A scaling factor of 79.429 was used to obtain the readout in fluorescence units, after iGEM Valencia_UPV 2018 [54]. The concentration profiles obtained from the model for the first toehold switch (hsa-miR-21–5p) is shown in Fig. 8(a). The trends of declining

miRNA and antimiR concentrations, and increasing complex concentration are evident. The complex formation occurred with exponential kinetics, and both the complex and the miRNA/antimiR attained saturation well before t = 500s. It is clear from Fig. 8(b) that the CTS and complex rapidly hybridized to open the CTS, with a significant drop in their concentrations concomitant while OTS concentration rises. The increase in OTS concentration permitted the translation of downstream GFP followed by the slow maturation kinetics of fluorescence. At about 500s, the complex was completely consumed, and OTS began to decay, while unhindered transcription allowed the CTS to continue to increase. Similar kinetics were observed in the case of the second toehold-switch (hsa-miR-20a-5p), with the smaller hybridisation rate constant slowing all intermediate reactions up to the final fluorescence readout (Supplementary File S7).

The concentration profiles for both the miRNA biomarkers informed

Table 5
Simulated fluorescence intensity variations with different concentrations of the complex over a 4 h period. The intensity shows sigmoidal kinetics with the logarithm of the trigger-complex concentration.

miRNA/anti-miR concentration (nM)	Complex (miR-antimiR-Toehold switch) (nM)	CTS (nM)	OTS (nM)	GFP (nM)	Mature GFP (μM)	Fluorescence Intensity
0.1	0.1	810	0.001	0.08	0.0055	4.76
1	0.98	810	0.01	0.07	0.055	47.73
10	9.87	810	0.13	0.79	0.551	477.17
25	24.69	810	0.34	1.97	1.38	1188.49
50	49.37	810	0.73	3.23	2.75	2382.83
75	74.05	810	1.13	5.58	4.32	3722.41
100	98.74	810	1.38	7.96	5.51	4754.40

(a)

(b)
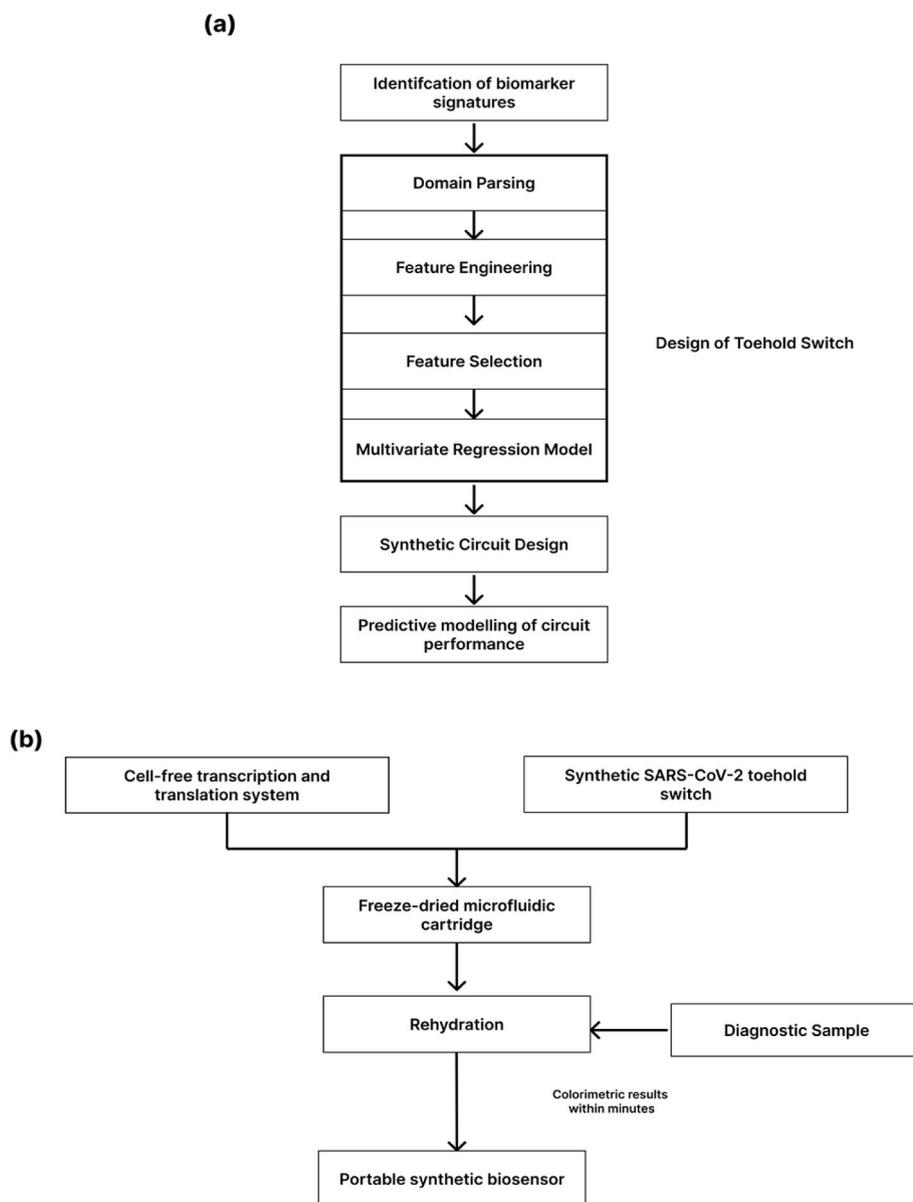
Fig. 10. (a) An end-to-end computational approach for design of toehold switch biosensors **(b)** A scheme for developing biosensor-based disease diagnostics; toehold switches and genomic signatures act as the biosensory element and analyte, respectively.

the range of complex concentration (viz. 100 pM to 10 μM) for modelling the fluorescence intensity emission. The plot of intensity against the complex concentration over a 2-h interval obtained a sigmoidal fit with a steep rise between 10 nM and 100 nM (Section S4, Supplementary File S7). This showed that the model is sensitive to changes in the concentration of the trigger miRNA biomarker in the range of interest, satisfying a necessary biosensor property [25,55–60]. A similar trend of sigmoidal kinetics was observed for the hsa-miR-20a-5p biomarker over a complex concentration range of 100 nM and 1 μM (Supplementary File S7). This suggested that the slower kinetics in the case of hsa-miR-20a-5p might require a higher biomarker concentration for detection. We repeated the calculations with a 4-h extended time interval, and found that the increased reaction time did not alter the sigmoidal kinetics, only yielding higher fluorescence intensity than the 2-h interval case ($R^2 = 0.99$) (Fig. 9). These theoretical calculations require experimental validation prior to clinical translation. Table 5 shows the concentration trajectories of various species and the predicted fluorescence intensity readout for different miRNA-antimiR complex concentrations over a 4-hr simulation.

(iii) Discussion

Towards the development of complete sensor circuits, medium-sized DNA constructs known as gBlocks (Integrated DNA Technologies; www.idtdna.com), housing one or more genes and conforming to the specifications in the Registry of Biological Parts [61], could be developed. In this direction, three independent gBlocks were designed using the standard BioBricks (https://biobricks.org), comprising the constitutive *E. coli* T7 promoter, a 5′ prefix sequence with EcoRI and XbaI recognition sites, and a 3′ suffix sequence with SpeI and PstI recognition sites, with the GFP-mut3b reporter gene. DNA sequences of the designed toehold-switches and genetic circuits are given in Supplementary File S8, while the gBlocks schematic, antimir design and BioBricks submissions are noted in Supplementary File S9. A limitation of the present study is that experimental validation with these synthetic circuits would be necessary before deployment in the clinic.

Neural networks have yielded useful applications for problems related to riboswitches [62]. Deep learning has been used for predicting toehold switch efficacy, with raw $R^2 \sim 0.43$–0.70 [63,64]. These models

are trained on a dataset of fused *cis*-triggers, significantly constraining their application to second-generation toehold switches. They require fixed-length inputs and do not lend to ready interpretation. Another application, MoiRNAifold is proprietary software and uses constraint-based linear programming to solve the inverse problem [65], and does not predict the efficacy of toehold switch structures.

In comparison, our work is an all-in-one package providing a light-weight solution to the problem of computational toehold design, efficacy prediction, and circuit modelling. Our tool is (i) agnostic of input sequence length, (ii) trained on free triggers, and (iii) capable of predicting in batch mode (any number of sequences in one go). Though the mechanistic understanding of RNA structure is yet incomplete, modelling attempts with interpretable features is a necessary step towards this goal. Further, miRNAs are an emerging class of biomarkers with desirable serum profiles, and typically induce a rogue stop codon in toehold switches immediately downstream to the start codon of the gene to be expressed. In this context, second-generation toehold switches are the viable solution, and it is hoped that our approaches would yield effective diagnostics for miRNA biomarkers of disease conditions. Any input sequence conforming to the generic toehold grammar will be processed by our rational pipeline thus:

(i) call to ViennaRNA RNAfold to parse the input sequence into its dot-bracket representation
(ii) call to GrammarParser to extract the segments of the toehold switch based on the dot-bracket representation
(iii) call to more ViennaRNA RNAfold utilities to obtain engineered feature values for use in the regression/neural network model
(iv) passing of feature values as arguments to the prediction script which yields the predicted dynamic range of the toehold switch sequence(s).

Geraldi et al. have highlighted the importance of synthetic biology in developing portable *in vitro* based diagnostic kits [66]. Toeholds embedded in genetic circuits enable the measurement of the intensity of expression of a downstream reporter protein – for e.g., GFP – that could yield a precise quantification of the trigger itself. Pardee has reviewed the potential of freeze-dried cell-free (FD-CF) systems in health care for diagnostic and sensing purposes, given their biosafe mode [67], and the promising opportunity for the rational design and manipulation of biological systems in relation to cell based systems has also been reviewed [68]. Fabricating microfluidic devices with our designed genetic circuits in a cell-free system followed by rigorous testing and validation would yield a point-of-care portable diagnostic tool for the real-time detection of early-stage cervical cancer.

Our workflow is modelled on the DBTL approach, which is analogous to the Design, Construct, Evaluate, Optimize (DCEO) approach in biotechnology [69], and Fig. 10 captures the methodology we have adopted here, in two parts. The end-to-end computational approach employed in this study is illustrated in Fig. 10(a). It includes four broad phases, namely identification of biomarker signatures, design of the biosensory toehold switch element, followed by the synthetic genetic circuit design comprising modular constructs and finally validating a proof-of-concept of the designed synthetic circuit using predictive modelling. Candidate toehold switches could be ranked for efficacy, and the predicted most efficacious toehold switch could be potentially used to optimize experimental workflows. This approach could be generalized to the design of portable, highly sensitive and specific biosensors for any condition, especially infectious diseases. The ongoing nCoV-SARS2 pandemic has highlighted the need for constant vigil for outbreak of infections. Such emerging agents require accurate, rapid, and scalable detection platforms, all requirements for which toehold biosensors are well-suited. Fig. 10(b) illustrates a scheme to produce such a toehold biosensor. Here the biomarker would be an optimal genomic fragment signature, unique to the pathogen but absent in the host. The toehold construct serves to signal the presence of the infectious agent that would then release the expression of the reporter gene, yielding a simple real-time visual readout. The proposed technique would generate a simple, reliable, rapid, portable and affordable printed diagnostic tool that could be made available at a global level in a point-of-care setting outside clinical diagnostic laboratory conditions. This biosafe product would be an asset in the field, especially in remote locations where it could help in curtailing the transmission and limiting the reproductive ratio of the infectious agent. Such a product would be vital in low-resource settings since it is sterile, abiotic and active for a year at room temperature. Extension of more than two biomarkers would allow for the simultaneous detection of, say, a family of coronaviruses, an imperative in our times. Thus, end-to-end scalable science is expected to overcome the limitations of ad-hoc diagnostic devices, thereby enabling affordable diagnosis in pandemic situations and other health emergencies.

4. Conclusion

Cervical cancer is a major public health issue with significant global burden of disease, but it is also an addressable one with the design of better molecular diagnostics. Early detection is necessary for effective treatment and greater compliance. In this work, three DBTL iterations were used to generate an *in silico* workflow for biomarker detection, sensor design and systems dynamics modelling. In the first cycle, we identified miRNAs that were significantly differentially expressed for early-stage cervical cancer using TCGA data and optimized down to two miRNA biomarkers based on prognostic significance. In the second cycle, we designed second-generation toehold switches and antimiRs for the two miRNAs. In this course, we developed a machine learning model of toehold-switch dynamic range that yielded adj.R^2 ~0.71, a major improvement in efforts in this direction. Finally we simulated the cell-free protein synthesis reactions to study the emergent kinetics. The results indicated that the fluorescence intensity underwent a sigmoidal transformation between 10 nM and 100 nM before reaching saturation. In summary, we have developed an end-to-end reproducible computational workflow for the sake of design of RNA biosensor devices for given conditions. Subsequent experimental validation would essentially democratize detection of emerging infectious diseases and life-threatening conditions with a ready transportable instrument based on real-time visual readout.

CRediT authorship contribution statement

Priyannth Ramasami S. Baabu: Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Shivaramakrishna Srinivasan:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Swetha Nagarajan:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Sangeetha Muthamilselvan:** Methodology, Validation, Formal analysis, Investigation, Visualization, Data curation, Visualization, Writing – original draft. **Thamarai Selvi:** Formal analysis, Software. **Raghavv R. Suresh:** Methodology, Validation, Formal analysis, Investigation, Visualization. **Ashok Palaniappan:** Conceptualization, Methodology, Software, Investigation, Formal analysis, Validation, Resources, Supervision, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests, financial or otherwise.

and infrastructural support to the iGEM 2019 project, from which this work originated (https://2019.igem.org/Team:SASTRA_Thanjavur). We would like to thank the anonymous reviewers for helping improve an earlier version of the manuscript. We would like to thank Ramit Bharanikumar and S. Prasanna Kumar for technical assistance. A.P. would like to acknowledge support from DST-SERB EMR/2017/000470.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2022.03.008.

References

[1] Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. Lancet 2007;370:890–907. https://doi.org/10.1016/S0140-6736(07)61416-0.

[2] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49. https://doi.org/10.3322/caac.21660.

[3] Stelzle D, Tanaka LF, Lee KK, Ibrahim Khalil A, Baussano I, Shah ASV, et al. Estimates of the global burden of cervical cancer associated with HIV. Lancet Global Health 2021;9:e161–9. https://doi.org/10.1016/S2214-109X(20)30459-9.

[4] Chandigarh, PGIMER. Health technology assessment of strategies for cervical cancer screening in India. 2019. https://dhr.gov.in/health-technology-assessment-strategies-cervical-cancer-screening-india. [Accessed May 2020]. accessed.

[5] Globocan. Cervix Uteri.https://gco.iarc.fr/today/data/factsheets/cancers/23-Cervix-uteri-fact-sheet.pdf; 2020[accessed January, 2021].

[6] Silverberg SG, Loffe OB. Pathology of cervical cancer. Cancer J 2003;9:335–47. https://doi.org/10.1097/00130404-200309000-00003.

[7] Ogedegbe G, Cassells AN, Robinson CM, DuHamel K, Tobin JN, Sox CH, et al. Perceptions of barriers and facilitators of cancer early detection among low-income minority women in community health centers. J Natl Med Assoc 2005;97:162–70. PMID: 15712779; PMCID: PMC2568778.

[8] Jearanaikoon P, Prakrankamanant P, Leelayuwat C, Wanram S, Limpaiboon T, Promptmas C. The evaluation of loop-mediated isothermal amplification-quartz crystal microbalance (LAMP-QCM) biosensor as a real-time measurement of HPV16 DNA. J Virol Methods 2016;229:8–11. https://doi.org/10.1016/j.jviromet.2015.12.005.

[9] Surriabre P, Torrico A, Vargas T, Ugarte F, Rodriguez P, Fontaine V. Assessment of a new low-cost, PCR-based strategy for high-risk human papillomavirus DNA detection for cervical cancer prevention. BMC Infect Dis 2019;19:842. https://doi.org/10.1186/s12879-019-4527-9.

[10] Wang Y, Chen M, Zhang L, Ding Y, Luo Y, Xu Q, et al. Rapid detection of human papilloma virus using a novel leaky surface acoustic wave peptide nucleic acid biosensor. Biosens Bioelectron 2009;24:3455–60. https://doi.org/10.1016/j.bios.2009.04.034.

[11] Roy S, Soh JH, Ying JY. A microarray platform for detecting disease-specific circulating miRNA in human serum. Biosens Bioelectron 2016;75:238–46. https://doi.org/10.1016/j.bios.2015.08.039.

[12] Pardini B, De Maria D, Francavilla A, Di Gaetano C, Ronco G, Naccarati A. MicroRNAs as markers of progression in cervical cancer: a systematic review. BMC Cancer 2018;18:696. https://doi.org/10.1186/s12885-018-4590-4.

[13] Filipów S, Ł Łaczmański. Blood circulating miRNAs as cancer biomarkers for diagnosis and surgical treatment response. Front Genet 2019;10:169. https://doi.org/10.3389/fgene.2019.00169.

[14] Cui M, Wang H, Yao X, Zhang D, Xie Y, Cui R, et al. Circulating MicroRNAs in cancer: potential and challenge. Front Genet 2019;10. https://doi.org/10.3389/fgene.2019.00626.

[15] Dembélé D, Kastner P. Fold change rank ordering statistics: a new method for detecting differentially expressed genes. BMC Bioinf 2014;15:14. https://doi.org/10.1186/1471-2105-15-14.

[16] van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6. https://doi.org/10.1038/415530a.

[17] Slomovic S, Pardee K, Collins JJ. Synthetic biology devices for in vitro and in vivo diagnostics. Proc Natl Acad Sci Unit States Am 2015;112:14429–35. https://doi.org/10.1073/pnas.1508521112.

[18] Vernon SD, Farkas DH, Unger ER, Chan V, Miller DL, Chen YP, et al. Bioelectronic DNA detection of human papillomaviruses using eSensor™: a model system for detection of multiple pathogens. BMC Infect Dis 2003;12. https://doi.org/10.1186/1471-2334-3-12.

[19] Ho NRY, Lim GS, Sundah NR, Lim D, Loh TP, Shao H. Visual and modular detection of pathogen nucleic acids with enzyme–DNA molecular complexes. Nat Commun 2018;9:3238. https://doi.org/10.1038/s41467-018-05733-0.

[20] Xu Y, Liu Y, Wu Y, Xia X, Liao Y, Li Q. Fluorescent probe-based lateral flow assay for multiplex nucleic acid detection. Anal Chem 2014;86:5611–4. https://doi.org/10.1021/ac5010458.

[21] Green AA, Silver PA, Collins JJ, Yin P. Toehold switches: de-novo-designed regulators of gene expression. Cell 2014;159:925–39. https://doi.org/10.1016/j.cell.2014.10.002.

[22] Pardee K, Green AA, Ferrante T, Cameron DE, DaleyKeyser A, Yin P, et al. Paper-based synthetic gene networks. Cell 2014;159:940–54. https://doi.org/10.1016/j.cell.2014.10.004.

[23] Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, Lee JW, et al. Rapid, low-cost detection of Zika virus using programmable biomolecular components. Cell 2016;165:1255–66. https://doi.org/10.1016/j.cell.2016.04.059.

[24] Hall RA, Macdonald J. Synthetic biology provides a toehold in the fight against Zika. Cell Host Microbe 2016;19:752–4. https://doi.org/10.1016/j.chom.2016.05.020.

[25] Takahashi MK, Tan X, Dy AJ, Braff D, Akana RT, Furuta Y, et al. A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. Nat Commun 2018;9:3347. https://doi.org/10.1038/s41467-018-05864-4.

[26] Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer Genome Atlas pan-cancer analysis project. Nat Genet 2013;45:1113–20. https://doi.org/10.1038/ng.2764.

[27] Broad Institute TCGA Genome Data Analysis Center. Analysis-ready standardized TCGA data from broad GDAC firehose 2016_01_28 run. Broad Inst MIT Harvard Dataset 2016. https://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/CESC/20160128dataset.

[28] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47. https://doi.org/10.1093/nar/gkv007.

[29] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

[30] Sarathi A, Palaniappan A. Novel significant stage-specific differentially expressed genes in hepatocellular carcinoma. BMC Cancer 2019;663. https://doi.org/10.1186/s12885-019-5838-3.

[31] Therneau T. A package for survival analysis in S. R package version. Survival (Lond). 2012.

[32] Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA package 2.0. Algorithm Mol Biol 2011;6:26. https://doi.org/10.1186/1748-7188-6-26.

[33] Green AA, Kim J, Ma D, Silver PA, Collins JJ, Yin P. Complex cellular logic computation using ribocomputing devices. Nature 2017;548:117–21. https://doi.org/10.1038/nature23271.

[34] To AC-Y, Chu DH-T, Wang AR, Li FC-Y, Chiu AW-O, Gao DY, et al. A comprehensive web tool for toehold switch design. Bioinformatics 2018;34:2862–4. https://doi.org/10.1093/bioinformatics/bty216.

[35] Chollet F. Deep learning with Python. second ed. Manning Publications, Shelter Island; 2021.

[36] CLSB-UK. Project BATMAN. CLSB-UK; 2017. http://2017.igem.org/Team.

[37] Stögbauer T, Windhager L, Zimmer R, Rädler JO. Experiment and mathematical modeling of gene expression dynamics in a cell-free system. Integr Biol 2012;4:494. https://doi.org/10.1039/c2ib00102k.

[38] Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. Software news and updates NUPACK: analysis and design of nucleic acid systems. J Comput Chem 2011;32:170–3. https://doi.org/10.1002/jcc.21596.

[39] Balacescu O, Balacescu L, Baldasici O, Tudoran O, Achimas-Cadariu P. The role of miRNAs in diagnosis, prognosis and treatment prediction in cervical cancer. Colposc. Cerv. Pathol. 2017. https://doi.org/10.5772/68011.

[40] Kang H-W, Wang F, Wei Q, Zhao Y-F, Liu M, Li X, et al. miR-20a promotes migration and invasion by regulating TNKS2 in human cervical cancer cells. FEBS Lett 2012;586:897–904. https://doi.org/10.1016/j.febslet.2012.02.020.

[41] Zhao S, Yao D, Chen J, Ding N, Ren F. MiR-20a promotes cervical cancer proliferation and metastasis in vitro and in vivo. PLoS One 2015;10:e0120905. https://doi.org/10.1371/journal.pone.0120905.

[42] Zamani S, Sohrabi A, Hosseini SM, Rahnamaye-Farzami M, Akbari A. Deregulation of miR-21 and miR-29a in cervical cancer related to HPV infection. MicroRNA 2019;8:110–5. https://doi.org/10.2174/2211536607666181017124349.

[43] Zhang Z, Wang J, Wang X, Song W, Shi Y, Zhang L. MicroRNA-21 promotes proliferation, migration, and invasion of cervical cancer through targeting TIMP3. Arch Gynecol Obstet 2018;297:433–42. https://doi.org/10.1007/s00404-017-4598-z.

[44] Wei W-F, Han L-F, Liu D, Wu L-F, Chen X-J, Yi H-Y, et al. Orthotopic xenograft mouse model of cervical cancer for studying the role of MicroRNA-21 in promoting lymph node metastasis. Int J Gynecol Cancer 2017;27:1587–95. https://doi.org/10.1097/IGC.0000000000001059.

[45] Cao M-M, Liu Q-X, Zou X, Fan X-C, Liu C, Zhang S-Y, et al. Identification of a serum three-microRNA signature for cervical cancer diagnosis. Chin Med J (Engl) 2021;134:1756–8. https://doi.org/10.1097/CM9.0000000000001327.

[46] Rao Q, Zhou H, Peng Y, Li J, Lin Z. Aberrant microRNA expression in human cervical carcinomas. Med Oncol 2012;29:1242–8. https://doi.org/10.1007/s12032-011-9830-2.

[47] Xin F, Liu P, Ma C-F. A circulating serum miRNA panel as early detection biomarkers of cervical intraepithelial neoplasia. Eur Rev Med Pharmacol Sci 2016;20:4846–51.

[48] Zhao S, Yao D, Chen J, Ding N. Circulating miRNA-20a and miRNA-203 for screening lymph node metastasis in early stage cervical cancer. Genet Test Mol Biomarkers 2013;17:631–6. https://doi.org/10.1089/gtmb.2013.0085.

[49] Chen J, Yao D, Li Y, Chen H, He C, Ding N, et al. Serum microRNA expression levels can predict lymph node metastasis in patients with early-stage cervical squamous cell carcinoma. Int J Mol Med 2013;32:557–67. https://doi.org/10.3892/ijmm.2013.1424.

[50] Liu SS, Chan KKL, Chu DKH, Wei TN, Lau LSK, Ngu SF, et al. Oncogenic microRNA signature for early diagnosis of cervical intraepithelial neoplasia and cancer. Mol Oncol 2018;12:2009. https://doi.org/10.1002/1878-0261.12383. –22.

[51] Wang J, Chen L. The role of miRNAs in the invasion and metastasis of cervical cancer. Biosci Rep 2019;39(3). https://doi.org/10.1042/BSR20181377.

[52] Quillet A, Saad C, Ferry G, Anouar Y, Vergne N, Lecroq T, et al. Improving bioinformatics prediction of microRNA targets by ranks aggregation. Front Genet 2020;10. https://doi.org/10.3389/fgene.2019.01330.

[53] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57. https://doi.org/10.1038/nprot.2008.211.

[54] Valencia_UPV. PRINTERIA. 2018. http://2018.igem.org/Team:Valencia_UPV. accessed October 2019.

[55] Lee K-H, Kim D-M. In Vitro use of cellular synthetic machinery for biosensing applications. Front Pharmacol 2019;10. https://doi.org/10.3389/fphar.2019.01166.

[56] Chan SK, Kuzuya A, Choong YS, Lim TS. DNA switch: toehold-mediated DNA isothermal amplification for dengue serotyping. SLAS Discov Adv Sci Drug Discov 2019;24:68–76. https://doi.org/10.1177/2472555218791743.

[57] Hwang Y, Kim SG, Jang S, Kim J, Jung GY. Signal amplification and optimization of riboswitch-based hybrid inputs by modular and titratable toehold switches. J Biol Eng 2021;15:11. https://doi.org/10.1186/s13036-021-00261-w.

[58] Pieters PA, Nathalia BL, van der Linden AJ, Yin P, Kim J, Huck WTS, et al. Cell-free characterization of coherent feed-forward loop-based synthetic genetic circuits. ACS Synth Biol 2021;10:1406–16. https://doi.org/10.1021/acssynbio.1c00024.

[59] Ma D, Shen L, Wu K, Diehnelt CW, Green AA. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. Synth Biol 2018;3. https://doi.org/10.1093/synbio/ysy018.

[60] Li H, Tang Y, Li B. Homogeneous and universal detection of various targets with a dual-step transduced toehold switch sensor. Chembiochem 2020;21:1418–22. https://doi.org/10.1002/cbic.201900749.

[61] International Genetically Engineered Machines. Registry of standard biological parts. http://parts.igem.org/Main_Pages. accessed May 2021.

[62] Premkumar KAR, Bharanikumar R, Palaniappan A. Riboflow: using deep learning to classify riboswitches with ~99% accuracy. Front Bioeng Biotechnol 2020;8. https://doi.org/10.3389/fbioe.2020.00808.

[63] Angenent-Mari NM, Garruss AS, Soenksen LR, Church G, Collins JJ. A deep learning approach to programmable RNA switches. Nat Commun 2020;11:5057. https://doi.org/10.1038/s41467-020-18677-1.

[64] Valeri JA, Collins KM, Ramesh P, Alcantar MA, Lepe BA, Lu TK, et al. Sequence-to-function deep learning frameworks for engineered riboregulators. Nat Commun 2020;11:5058. https://doi.org/10.1038/s41467-020-18676-2.

[65] Minuesa G, Alsina C, Garcia-Martin JA, Oliveros JC, Dotu I. MoiRNAiFold: a novel tool for complex in silico RNA design. Nucleic Acids Res 2021;49:4934–43. https://doi.org/10.1093/nar/gkab331.

[66] Geraldi A, Giri-Rachman EA. Synthetic biology-based portable in vitro diagnostic platforms. Alexandria J Med 2018;54:423–8. https://doi.org/10.1016/j.ajme.2018.11.003.

[67] Pardee K. Perspective: solidifying the impact of cell-free synthetic biology through lyophilization. Biochem Eng J 2018;138:91–7. https://doi.org/10.1016/j.bej.2018.07.008.

[68] Tinafar A, Jaenes K, Pardee K. Synthetic biology goes cell-free. BMC Biol 2019;17:64. https://doi.org/10.1186/s12915-019-0685-x.

[69] Chen X, Gao C, Guo L, Hu G, Luo Q, Liu J, et al. DCEO biotechnology: tools to design, construct, evaluate, and optimize the metabolic pathway for biosynthesis of chemicals. Chem Rev 2018;118:4–72. https://doi.org/10.1021/acs.chemrev.6b00804.