
# BUILDING A PARALLEL DECISION-MAKING SYSTEM BASED ON RULE-BASED CLASSIFIERS IN MOLECULAR ROBOTICS

Wibowo Adi[1,2] and Kosuke Sekiyama[1]

[1]Department Micro-Nano System Engineering, Nagoya University, Nagoya, Japan

[2]Department of Informatics, Diponegoro University, Semarang, Central Java, Indonesia

Email: wibowo@robo.mein.nagoya-u.ac.jp

***Submitted:*** *Feb. 20, 2015* ***Accepted: Mar. 27, 2015*** ***Published: June 1, 2015***

***Abstract.*** Controlled drug delivery based on cellular components can be achieved by exploiting disease-specific properties, but these require a rapid, sensitive, and selective method of detection in a biomolecular system. We propose a parallel decision-making system for disease detection and classification based on the fact that DNA computing along with biomolecular systems can be subjected to massively parallel processing. We designed and programmed a DNA strand displacement reaction to implement rule-based classifiers from a binary tree classification as a decision-making system. In our framework for molecular robot development, the system components of molecular robots and simple classifier rules were used to alleviate the computational burden. The design consists of a basic model that generates rule-based classifier gates in several binary tree and cancer classifications based on micro (mi)RNA expression. Simulation results showed that detection and classification were rapid using this system. Moreover, experiments using the synthetic miRNA hsa-miR-21 demonstrated that our model could be a feasible decision-making system for drug delivery.

**Index terms:** Molecular robotics, DNA strand displacement, Rule-based classifiers, Binary tree classification.

# I. INTRODUCTION

The development of intelligent drug devices for disease diagnosis and therapy has gained increasing attention in recent years. A drug delivery system must sense the biological environment (biosensing) and restrict the release of a drug to specific areas of disease [1]. Micro (mi)RNAs are potential bio-sensing targets given that they have been implicated in many diseases, including cancer [2]; recent studies have also described the miRNA profiles of various allergic inflammatory diseases, including asthma, eosinophilic esophagitis, allergic rhinitis, and atopic dermatitis [3]. In addition, specific miRNAs have been identified as regulators of pathogenic mechanisms in allergic inflammation [2, 3], such as the polarization of adaptive immune responses and T cell activation [4], while another study found distinct patterns of miRNA expression in lung cancer, colorectal cancer, and diabetes, providing evidence that miRNA fingerprints are present in many diseases [5].

Nano-scale robots (nanorobots [6]) that can interact with a biomolecular system may be useful for miRNA detection. Nanorobots must have at least some of the following functions: actuation, sensing, signaling, classification, decision-making, and swarm behavior at the nano scale [7]. DNA is the perfect biomolecule for creating nanorobots given its size and capacity for hydrogen bonding between complementary bases that allows complex shapes and structures as well as DNA-RNA hybrids to be generated at the nano scale [8]. In addition, DNA is considered as an autonomous molecule that does not require human control [9]. Based on these advantages, DNA is the main candidate material for molecular robotics [10]. However, molecular robots used for programmed and controlled drug delivery require a decision-making mechanism that interacts with disease-specific biomolecules.

DNA computing and molecular programming are the main tools used to develop decision-making systems based on chemical features of DNA [11], RNA [12], gel [13], and enzymes [14]. Previous studies have proposed computing processes that exploit DNA strand displacement (DSD) [15], such as logic gates [16], amplification [17], and logic circuits or applied neural network computation [18] from seesaw gates [19]. However, systems that are capable of classifying biomolecules require complex molecular computation processes that are not possible for molecular robots.

To address the issue of computational complexity for biomolecule classification, we propose the massively parallel processing of simple classification algorithms in a synthetic DSD reaction. The massively parallel processing capabilities of DNA computers [11] can provide simultaneous decision-making and can potentially accelerate processes in a large system. To obtain a simple classification algorithm in the DSD reaction, rule-based classifiers from a binary tree classification were used; these generated descriptive models that were easy to interpret but have comparable performance to decision tree classifiers [20]. The ability to explain the reason(s) for a decision is crucial, and binary tree-based classification of diseases such as cancer has been proposed by other investigators [21, 22]. Most studies suggest that miRNA expression is a feature of cancer classification, diagnosis, and disease progression. Using a disease-based classification tree, we designed a DSD reaction of rules and classifier algorithms. A decision-making system can therefore be integrated into a molecular robot designed to function as a drug delivery system. The framework for molecular robot development is presented in Section II. In Section III, we describe the proposed model; in Sections IV and V, results of the simulation and experiment, respectively, are presented. Conclusions and future directions are discussed in Section VI.

## II. FRAMEWORK

The framework for molecular robot development was inspired by previous studies in nucleic acid nanotechnology, which can be broadly divided into structural, dynamic, and interface DNA nanotechnology. Structural DNA nanotechnology concerns the self-assembly of nucleic acid structures with well-defined shapes, sizes, and/or patterns. Dynamic DNA nanotechnology pertains to non-equilibrium systems in which DNA molecules undergo a series of conformational changes that physically or chemically modify the environment. Interface DNA nanotechnology uses nucleic acids as a tool for controlling other nanoscale materials such as carbon nanotubes and gold nanoparticles [23]. The framework for developing a molecular robot that functions as a drug delivery system is illustrated in Figure 1.

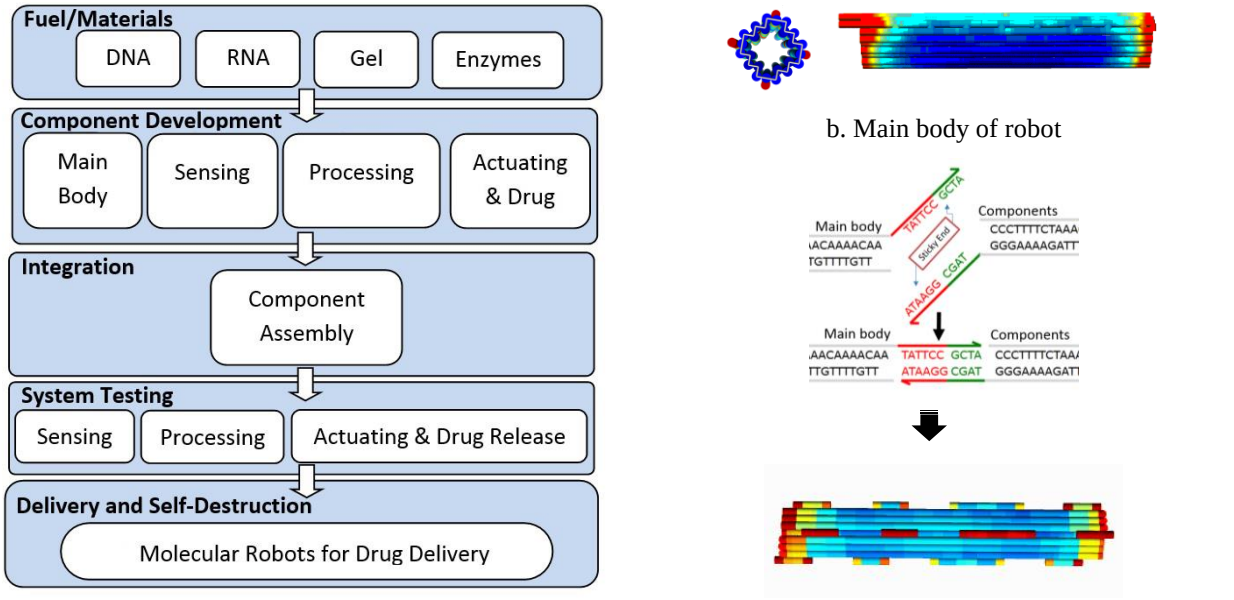



**Figure 1.** Framework for molecular robot development.

Molecular robots are structured from DNA, RNA, gels, and enzymes and include five components: the main body, and sensing, processing, actuating, and drug release components (Fig. 1a). Constructing the main body of the robot depends on structural DNA nanotechnology required to synthesize and assemble nucleic acid complexes and materials with a static, equilibrated endpoint. Various structures have been developed using the DNA origami method [24]—including two- [25] and three-dimensional [26] as well as discrete structures [27]—in which a long, single-stranded DNA scaffold is deformed into the desired body shape using glue strands or staples. In the present study, the main body of the capsule/tube was designed with caDNAno software [28], while CanDo software [29] was used to model the stability of the DNA origami structure (Fig. 1b).

The sensing component is responsible for target recognition and signal transduction. The target can be any chemical or biological material such as small organic molecules, peptides, proteins, nucleic acids, carbohydrates, or even whole cells [30]. Signal transduction converts molecular recognition into physically detectable signals such as single-stranded DNA, fluorescence, color, electrochemical signals, or magnetic resonance changes [30].

The processing component can be viewed as a computational device that processes physical or chemical inputs to generate an output based on a set of logical operators [31]. This

component obtains information obtained from the molecular environment, which is interpreted by DNA computing or molecular programming to determine the necessary response [32]. The actuating and drug release components function as the muscles of the robot by converting stored energy into movement. The possibility of using DNA as an actuator was first suggested with autonomous DNA walkers [33].

Each component is connected to the body by its own interface. This step employs the self-assembly model, sticky-end cohesion [34], and the strand displacement model to integrate each component with the body. The integration of the main body and components was simulated with CanDo software (Fig. 1c). Before delivery to the target by the robot, the system will be tested by evaluating each function and measuring effectiveness and durability.

## III. PARALLEL DECISION-MAKING SYSTEM

This section focuses on sensing and processing as steps towards decision making. We first present rule-based classifiers in a binary tree classification, followed by the basic principles of DSD. We then describe the implementation of the parallel decision-making system based on rule-based classifiers in DSD.

### *A. Rule-based classifiers in a binary tree classification*

Classification is an important problem in decision making that requires a systematic approach to building classification models from an input dataset. Examples include decision tree, rule-based, and naïve Bayes classifiers, neural networks, and support vector machines. Each technique employs a learning algorithm to identify a model that best fits the relationship between the attribute set (input) and class label (output) of the input data. The model generated by the learning algorithm should closely fit the input data as well as correctly predict the class labels of novel records [20]. We used rule-based classifiers in binary trees where leaves and branches represent classifications and feature-based splits leading to classifications, respectively [35].

Rule-based classifiers depend on a set of IF-THEN rules [20]; those of the model were represented in a disjunctive normal form, $R = (r_1 \vee r_2 \vee \ldots r_k)$, where $R$ is the rule set and each $r_i$ is a classification rule or disjunct. Thus, each classification rule was expressed in the following form:

$$r_i: IF\ (condition_i)\ THEN\ y_i \quad (1a)$$

where IF (or left side of the rule) was the rule antecedent or precondition containing a conjunction of attribute tests, and

$$Condition_i = (A_1\ op\ \mu_1) \wedge (A_2\ op\ \mu_2) \wedge \ldots (A_k\ op\ \mu_k), \quad (1b)$$

where $(A_j\ op\ \mu_j)$ was the attribute-value pair and *op* was a logical operator chosen from the set {=, ≠, <, >, ≤, ≥}. Each attribute test $(A_j\ op\ \mu_j)$ was known as a conjunct. THEN (or the right-hand side) was the rule consequent or conclusion containing the predicted class $y_i$ [20].

A rule *r* covered record *x* if the precondition of *r* matched the attributes of *x*; *r* was also said to be fired or triggered when it covered a given record. That is, if the condition of the rule was satisfied, the conclusion was inferred or deduced. The rules were used for classification if their consequents were a predefined class and their antecedents/preconditions contained conditions of various features and their corresponding values [36].

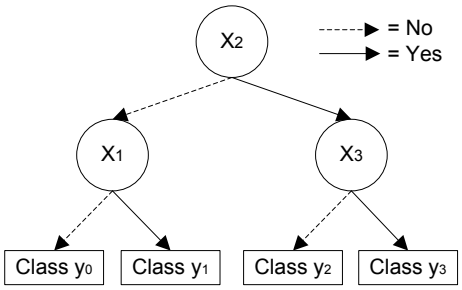



**Figure 2.** Fully weighted binary tree of three input fields and four classes

A binary tree comprised three sets *L*, *S*, and *R*, where *L* and *R* were binary trees (or are empty), and *S* was an individual set. The single element of *S* was the root, while *L* and *R* were the left and right subtrees of the root, respectively [37]. A sample binary tree is shown in Figure 2, in which the root is $X_2$, and the left right subtrees are $X_1$ and $X_3$, respectively. The basic concerns in designing a binary tree classification are the separation of two groups of classes among training samples at each non-terminal node ($X_1$ and $X_3$), where decision-making is necessary, and the

choice of the subset of features which is most effective in separating these two groups of classes [38].

We assumed that this binary tree was already optimized based on the above criteria. In our proposed method, the tree was formed as a fully weighted binary tree. We assumed that the left and right leaves were negative and positive criteria, respectively. The discovery of decision rules for forming branches or segments beneath the root node was based on a method that extracts the relationship between the object of analysis—which serves as the target field in the data—and one or more fields that serve as input fields for creating branches or segments. The values in the input field were used to estimate the likely value in the target field, also referred to as an outcome, response, or class [37]. To extract rules from the decision tree, one rule was created for each path from the root to a leaf node. Each splitting criterion along a given path was logically AND from the rule antecedent (IF). The leaf node held the class prediction, forming the rule consequent (THEN) [38]. The rules ($r_0 \ldots r_3$) extracted from Figure 2 were as follows:

$$r_0: \text{IF } (X_1 = No) \wedge (X_2 = No) \wedge (X_3 = No) \text{ THEN } class\ y_0 \quad (1c)$$

$$r_1: \text{IF } (X_1 = Yes) \wedge (X_2 = No) \wedge (X_3 = No) \text{ THEN } class\ y_1 \quad (1d)$$

$$r_2: \text{IF } (X_1 = No) \wedge (X_2 = Yes) \wedge (X_3 = No) \text{ THEN } class\ y_2 \quad (1e)$$

$$r_3: \text{IF } (X_1 = No) \wedge (X_2 = Yes) \wedge (X_3 = Yes) \text{ THEN } class\ y_3 \quad (1f)$$

*B. DSD*

Strand displacement is the process by which two strands with partial or full complementarity hybridize to each other, displacing one (referred to as the displacement domain, which is around 20 nucleotides [nt]) or more pre-hybridized strands in the process. Strand displacement can be initiated at complementary single-stranded domains—referred to as toeholds—of 4–9 nt and progresses through a branch migration process resembling a random walk [15, 39]. Figure 3 illustrates the strand displacement reaction, which shows displacement L as a long domain holding from the left side of toehold domain s (s^ and s^* are complementary, as are L and L*). The resultant reaction was the displacement of single strand L. In the system design, we used only the domain name without including details of nucleotide sequence.

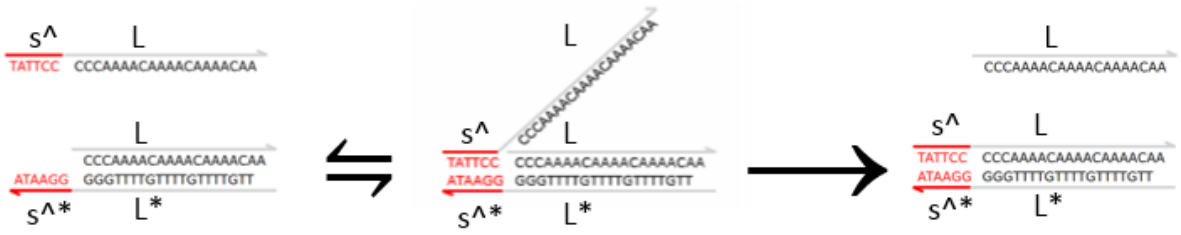



**Figure 3.** Strand displacement reaction.

*C. Implementation of a parallel decision-making system based on rule-based classifiers in strand displacement*

We assumed that the optimal design of binary tree classification that achieved the highest possible accuracy could be determined computationally. Thus, the smallest DSD reaction was used in order to reduce computational complexity. Decision-making rules were selected at each non-terminal node of the binary tree classification. To build a rule-based classifier, a set of rules identifying key relationships between the attributes of the dataset (input fields) and the class label was needed. The optimal IF-THEN rules were extracted directly from the binary tree classification. The list of rules for classification was implemented in DNA strand reactions and parallel processes occurring in the classification reaction. The many classification rules in rule set *R* provided exhaustive coverage, including the rules shown in equations (1e) and (1f), both of which were triggered by the same record and had the same combination of attribute values. Further, the same combination of attribute values require a complex logic gates or circuits to get decision-making system. To address strand displacement and simplify reactions, our proposed model did not include complex logic gates or circuits. This approach allowed a test record to trigger multiple classification rules in parallel. There were two kinds of parallel systems in our proposed model; that is, sensing and encoding systems of IF-THEN rules.

The parallel sensing system as the rule antecedent/precondition addressed input signals from the outside by receiving input criteria from environmental conditions. Numerous input signals were detected simultaneously without dependence on other types of input. Figure 4 shows the parallel sensing reactions of three input fields as the rule antecedent/precondition based on the binary tree shown in Figure 2. We demonstrated that the parallel sensing gates contained a domain of binary tree classification input fields ($X_1$, $X_2$, and $X_3$) that were partially split into two

domains, i.e., toehold (colored; $x_1$^, $x_2$^, $x_3$^) and displacement ($x_1$, $x_2$, and $x_3$). The input field signal strands (<$x_1$^ $x_1$>, <$x_2$^ $x_2$>, and <$x_3$^ $x_3$>) beside the gates hybridized to bound complementary domains on the gates via uncovered toeholds ($x_1$^*, $x_2$^*, and $x_3$^*), resulting in branch migration through recognition domains. The previously bound strands were predicted to fall off given that they were attached to gate base strands by a short toehold ($y_1$^, $y_2$^, and $y_3$^). The now-bound signal would have an uncovered toehold on the other side; therefore, the now-free signals (<$x_1$ $y_1$^ *1*>, <$x_2$ $y_2$^ *2*>, and <$x_3$ $y_1$^ *1*>) were able to reverse the process symmetrically. To obtain a decision class from each input, the now-free signals containing an encoded number as the domain output ($2^0$, $2^1$, $2^2$ …) for each input were used as the sensing output and decided the class position based on IF-THEN rules. The class numbering was only applicable to binary tree schema. Moreover, it was supposed that the IF-THEN rule detected a positive result if it was overexpressed during the sensing process.

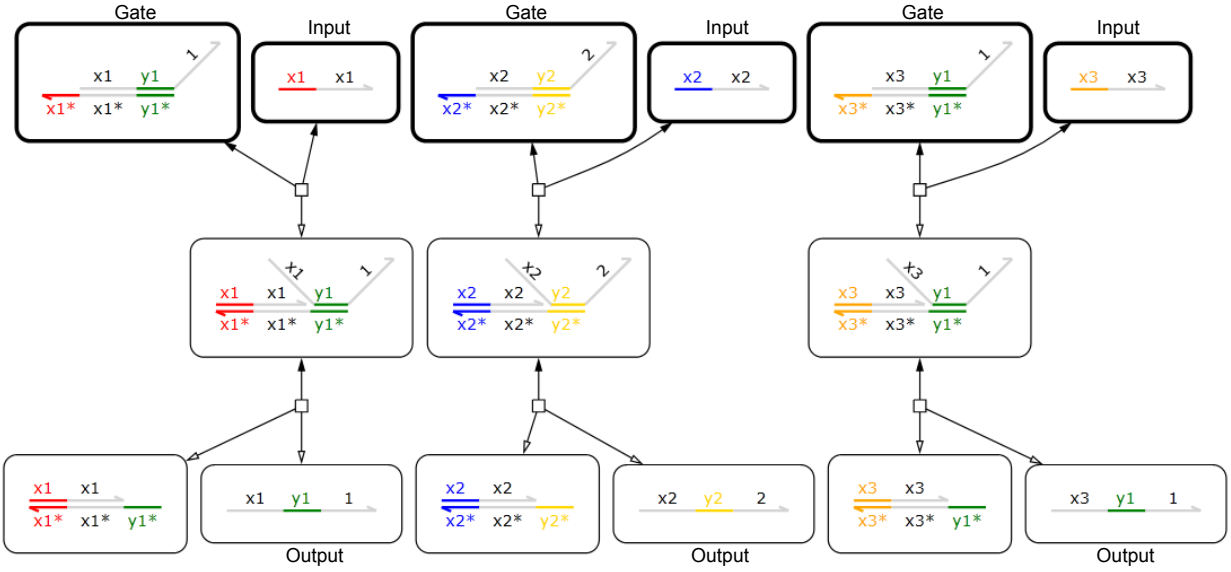



**Figure 4.** Parallel sensing reactions of three input fields as the rule antecedent or precondition.

The parallel sensing system was designed to sense DNA as well as miRNA input. We therefore selected miRNA and DNA sequences that can undergo strand displacement. For example, miRNA hsa-miR-21 has a 22-base pair sequence UAGCUUAUCAGACUGAUGUUGA that is split into two domains: a toehold or miR21t domain from nucleotides 1–6 (UAGCUU) and a displacement or miR21d domain from

nucleotides 7–22 (AUCAGACUGAUGUUGA). The strand displacement process of hsa-miR-21 is shown in Figure 5.

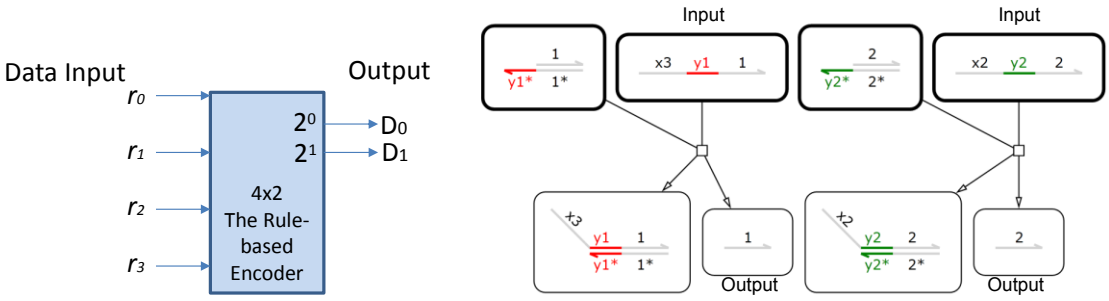



**Figure 5.** Strand displacement of miRNA hsa-miR-21.

The second reaction was the parallel encoding system as the rule consequent, which was used to show decision results in simulations or experiments and break reverse the process of the sensing system. We created a binary encoding of classes to simplify recognition. Classification rules and class numbers encoded as two-digit binary numbers based on the binary tree in Figure 2 are shown in Table 1. These values were used to generate transducer-encoding signals and obtain a decision result from the parallel sensing system, as shown in Figure 6, which illustrates rule/class encoding and the reaction process of the encoding signal. The now-free signals from sensing gates bound to encoding gates via a short toehold and produced only inert waste molecules without toeholds. In experiments, the displacement domain in the encoding system was marked by a fluorescent signal; as such, classification results were detectable by fluorescence spectrometry, and the fluorescent class signal could be observed without DNA/RNA purification.

**Table 1.** Classification rules and encoding of the four-class binary tree classification

| | | Input | | | Output | | Class |
|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $X_3$ | $D_1$ | $D_0$ | |
| Rules | $r_0$ | 0 | 0 | 0 | 0 | 0 | $y_0$ |
| | $r_1$ | 1 | 0 | 0 | 0 | 1 | $y_1$ |
| | $r_2$ | 0 | 1 | 0 | 1 | 0 | $y_2$ |
| | $r_3$ | 0 | 1 | 1 | 0 | 0 | $y_3$ |

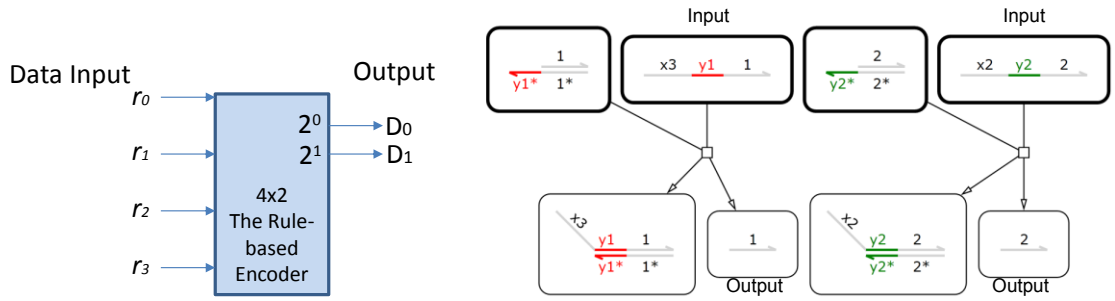



**Figure 6.** DNA reaction in the parallel encoding system.

In the proposed model, the sensing and encoding systems were combined in the same gates, which was akin to the tethering of DNA molecules [40], to increase their local concentration as well as reaction rates [40] and thereby increase the speed of the parallel reaction between both systems in experiments. We demonstrated how the implementation of an interface at this gate, termed the chromosome, could be integrated with the body of the molecular robot (Fig. 7). Additionally, a full weight binary tree classification of eight classes revealed the design pattern of the chromosome (Fig. 8). The classification rules of the eight classes and class numbers encoded as binary numbers in the tree are shown in Table 2.

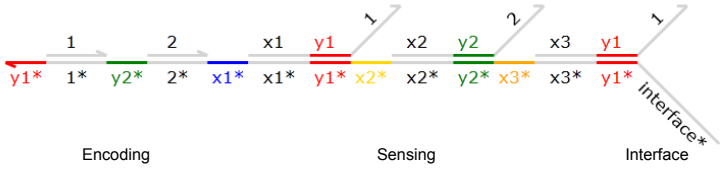



**Figure 7.** System and interface chromosomes.

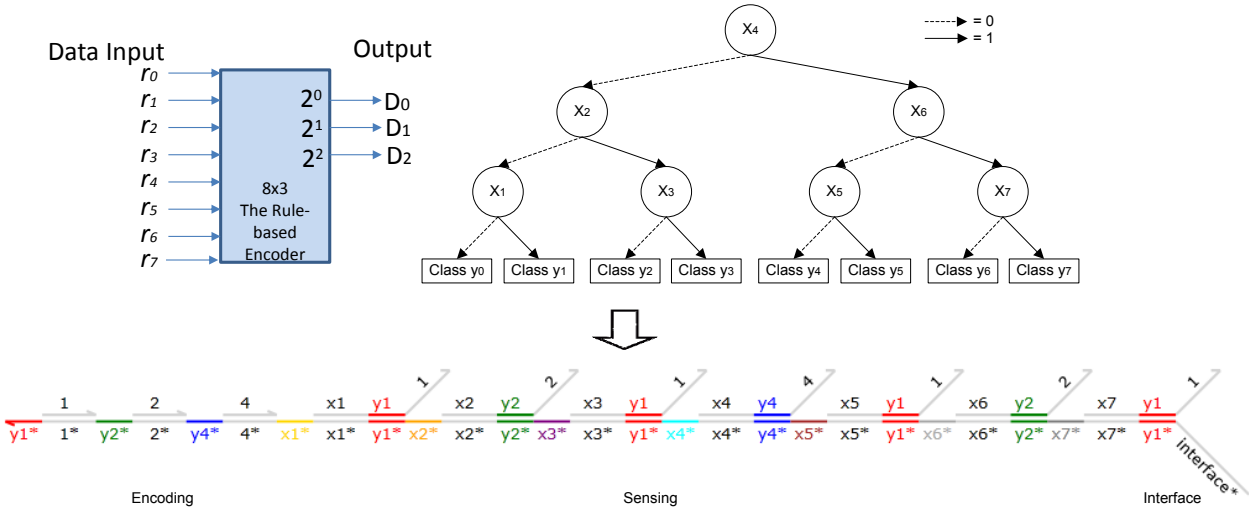



**Figure 8.** Full weight binary tree classification of eight classes and the chromosome.

**Table 2.** Classification rules and encoding of the eight-class binary tree classification

| | | Input | | | | | | | Output | | | Class |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $D_2$ | $D_1$ | $D_0$ | |
| Rules | $r_0$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $y_0$ |
| | $r_1$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | $y_1$ |
| | $r_2$ | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | $y_2$ |
| | $r_3$ | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | $y_3$ |
| | $r_4$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | $y_4$ |
| | $r_5$ | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | $y_5$ |
| | $r_6$ | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | $y_6$ |
| | $r_7$ | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | $y_7$ |

## VI. SIMULATION

We used visual DSD [41] to simulate the displacement process and implemented parallel sensing and encoding systems for decision-making. The chemical reaction network of the decision-making system in one chromosome ($C$) is shown in the following equation.

$$Input + C \leftrightarrow C' \rightarrow C'' + Output \quad (2)$$

Figure 9 shows an example of the parallel process reaction cascade of DSD in a chromosome of eight classes based on equation (2). This process was simulated by class $y_7$, which had the criteria

$X_4$, $X_6$, and $X_7$ as an input condition. The output of this reaction consisted of signals 1, 2, and 4 (shown in different colors). The parallel process showed that the sensing and transducing systems could operate simultaneously using one gate. Figure 10 shows the simulation results of a signaling class using three different color signals to represent the output of a decision or classification result. These results demonstrate that the gate can assign input to a specific class, as shown in Table 2.

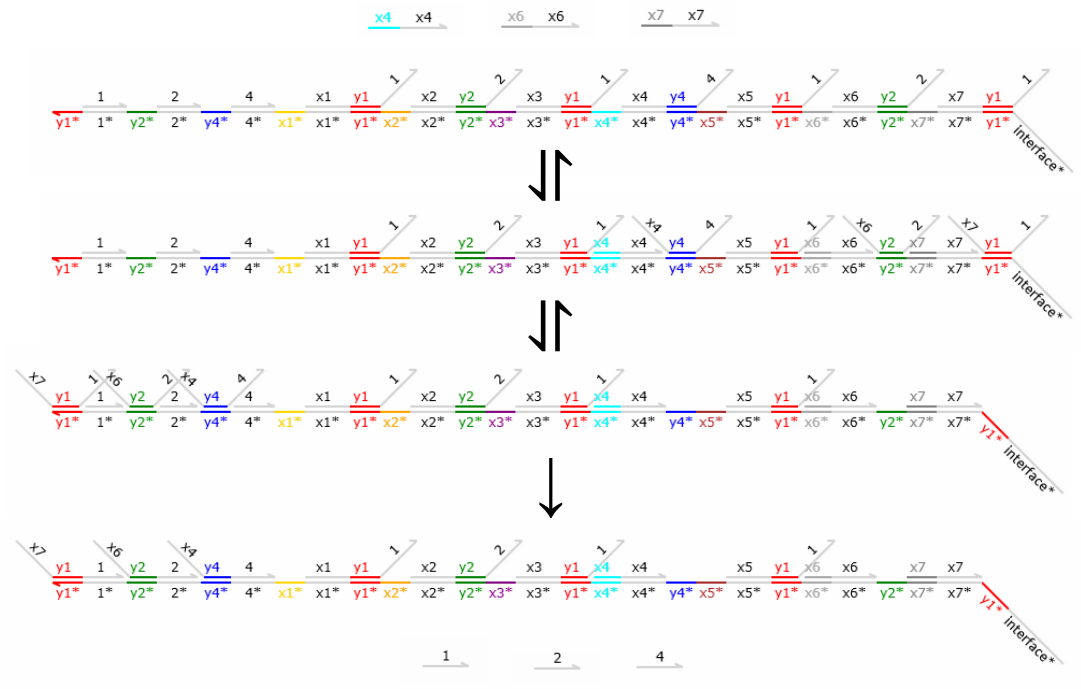
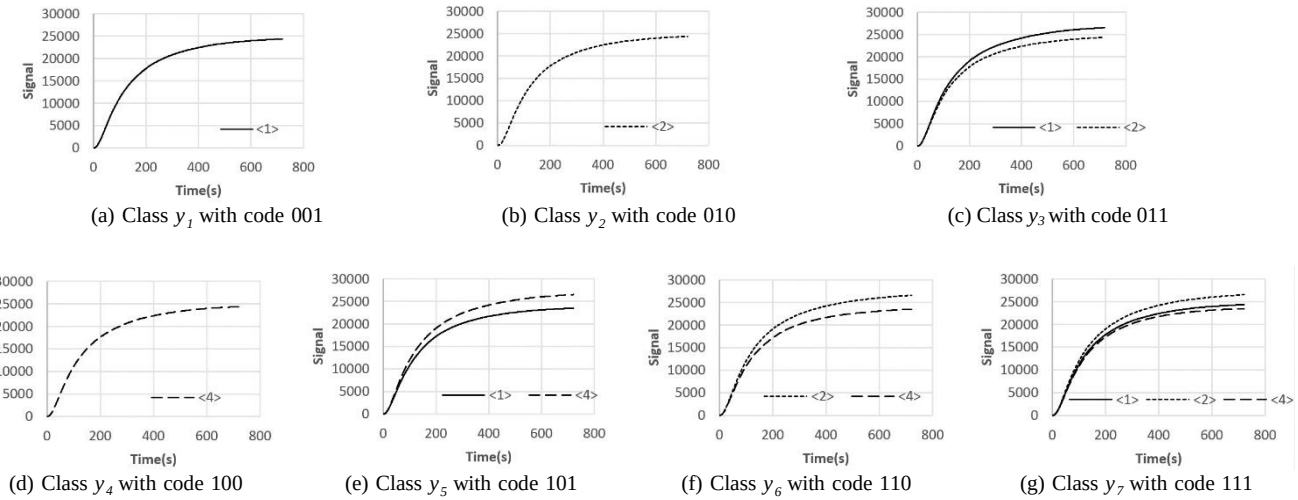


**Figure 9.** Examples of states and cascade reactions of chromosomes.




(a) Class $y_1$ with code 001 (b) Class $y_2$ with code 010 (c) Class $y_3$ with code 011

(d) Class $y_4$ with code 100 (e) Class $y_5$ with code 101 (f) Class $y_6$ with code 110 (g) Class $y_7$ with code 111

**Figure 10.** Simulation of signaling classes.

The proposed decision-making system was designed to be integrated into the molecular robot for drug delivery; as such, it had to incorporate miRNA detection and classification. Since visual DSD can only simulate DNA displacement, we assumed that a miRNA signal was used as an input in the simulation. The classification of cancer was thus simulated based on the reports [21]. Figure 11 illustrates the binary tree classification of cancer. The rules in the classification diagram of the original report—which used the left branch as positive criteria—were reversed; that is, in the present study the left branch was used as negative criteria. Numbered miRNAs used as nodes in the decision tree classifier are shown in Table 2 of the original report [21]. Classes of cancer were represented by the letters $y_1$ to $y_{15}$. The set of rules for the binary tree classification of cancer were as follows.

*$r_1$: IF miRNA group 7, THEN class $y_1$: Meninges*
*$r_2$: IF miRNA group 6, THEN class $y_2$: Brain*
*$r_4$: IF miRNA group 4, THEN class $y_4$: Melanocytes*
*$r_6$: IF miRNA group 4 ∧ 5 THEN class $y_6$: Lymph node*
*$r_{10}$: IF miRNA group 3 ∧ 16 THEN class $y_{10:}$ Prostate*
*$r_{11}$: IF miRNA group 3 ∧ group 16 ∧ group 17 THEN class $y_{11}$: Breast*
*$r_{12}$: IF miRNA group 3 ∧ group 12 THEN class $y_{12}$: Lung (carcinoid)*
*$r_{15}$: IF miRNA group 3 ∧ group 12 ∧ group 13 ∧ group 14 THEN class $y_{15}$: Colon*

Class number was encoded as a four-digit binary number (Table 3). Based on the tree and encoding table, a chromosome gate for cancer classification was designed (Fig. 12).

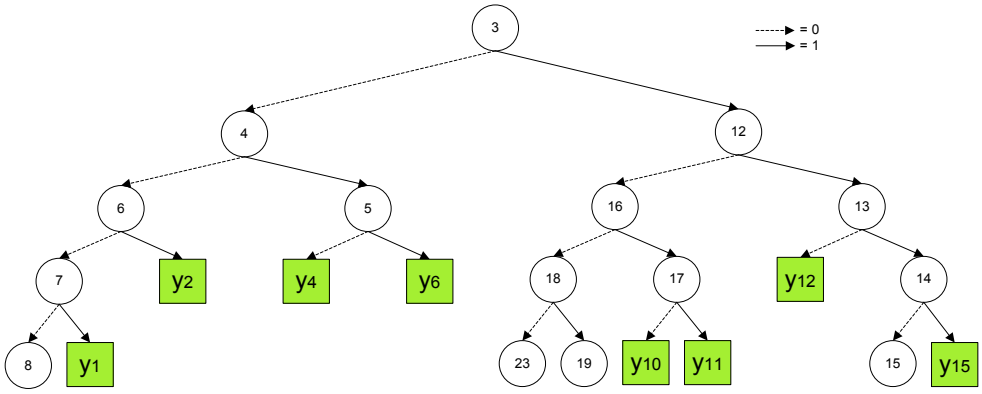



**Figure 11.** Illustration of the classifier structure for the cancer decision-making tree [21] (levels 3 to 8 only).

**Table 3.** Encoding of cancer classes

| Class | Cancer | Binary Encoding |
|---|---|---|
| $y_1$ | Meninges | 0001 |
| $y_2$ | Brain | 0010 |
| $y_4$ | Melanocytes | 0100 |
| $y_6$ | Lymph node | 0110 |
| $y_{10}$ | Prostate | 1010 |
| $y_{11}$ | Breast | 1011 |
| $y_{12}$ | Lung (carcinoid) | 1100 |
| $y_{15}$ | Colon | 1111 |

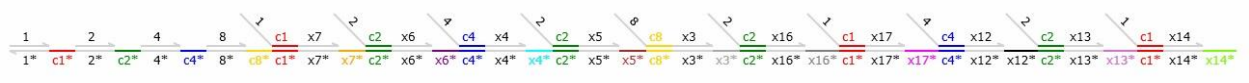
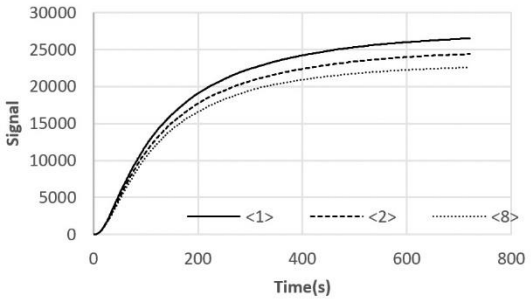
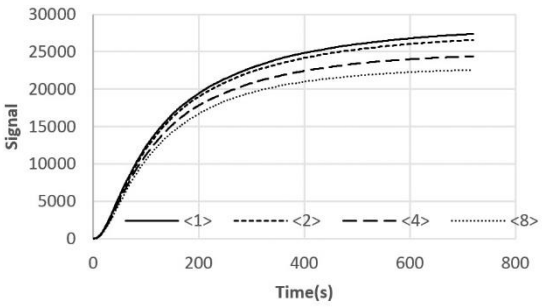


**Figure 12.** Chromosome gate for cancer classification.

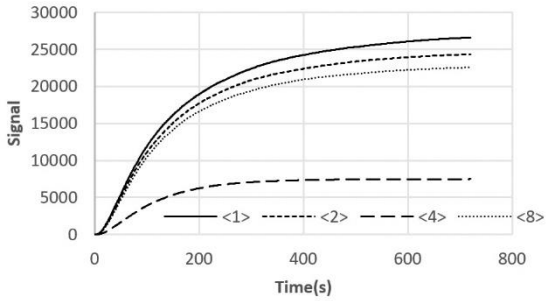
For simulation conditions, the miRNA input surrounding the gates was set according to specific criteria. For example, if the input was miRNA groups 3, 16, and 17, then the proposed gates would provide a clear output signal for the breast cancer class (Fig. 13a). For miRNA input from groups 3, 12, 13, and 14, the gates would produce output signals for colon cancer (Fig. 13b). However, some miRNA input signals are present in every cancer cell (see Supplemental Table 2 in ref. [21]), which can lead to an abnormal output signal; for instance, in the example shown in Figure 13c, the level of signal 4 is lower than that of other signals. Thus, a threshold system was applied to each encoding signal. Assuming that signal 4 had not reached the minimum limit of the threshold, the signals in Figure 13c were assigned to the breast cancer class but with noise.




(a) Class $C_{11}$ or breast cancer (code 1011)




(b) Class $C_{15}$ or colon cancer (code 1111)




(c) Class $C_{11}$ or breast cancer (with noise)

**Figure 13.** Simulation of signaling cancer class

## V. EXPERIMENT

The miRNA hsa-miR-21, which is overexpressed in many cancer types [42], was selected as an input for testing the sensing and encoding system. Gates were developed for hsa-miR-21. The cascade reaction of the experiment is shown in Figure 14. Displacement of a quencher and fluorophore duplex led to a fluorescent output at room temperature, which was used in domain 1.

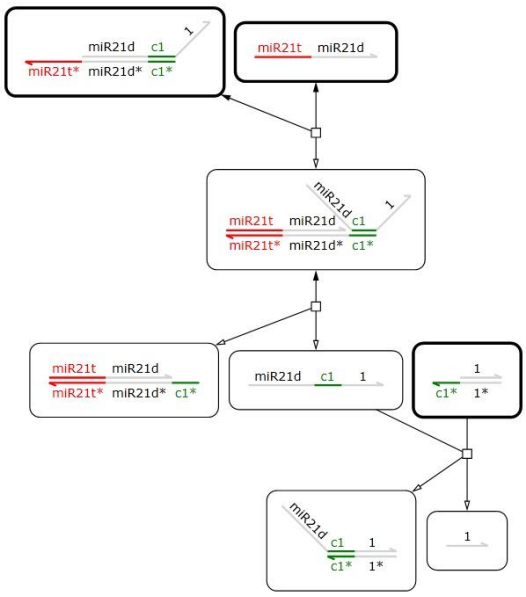



**Figure 14.** Strand displacement for miRNA hsa-miR-21.

The MISSION miRNA hsa-miR-21 mimic was purchased from Sigma-Aldrich (St. Louis, MO, USA) and diluted in sterile water. All DNA oligonucleotides used in this experiment were from Operon (Tokyo, Japan). To avoid mismatches and secondary structure formation, DNA sequences were optimized using Nupack nucleic acid software [43]. The DNA was purified by high-performance liquid chromatography and coding strands were modified with the ROX fluorophore and Black Hole Quencher 2. The DNA was resuspended as a 50 µM stock solution in distilled water and stored at −20 °C until use. DNA and miRNA sequences used in this study are shown in Table 4.

**Table 4.** DNA and miRNA sequences

| Strand | Sequence* |
|---|---|
| MiRNA hsa-miR-21 | UAGCUUAUCAGACUGAUGUUGA |
| Sensing – upper | ATCAGACTGATGTTGATACCAACCAATTTCTAACCTAAACAA |
| Sensing – lower | TTGGTATCAACATCAGTCTGATAAGCTA |
| Encoding – upper | CCAATTTCTAACCTAAACAA[BHQ2a-Q] |
| Encoding – lower | [AminoC6+ROX]TTGTTTAGGTTAGAAATTGGTTGGTA |

*Toehold domains are underlined.

BHQ2a-Q, Black Hole Quencher 2.

Sensing and encoding gates were prepared from DNA stock solutions by mixing the two strands (2~4 μM each) in 1× TAE/$Mg^{2+}$ buffer composed of 40 mM Tris (pH 8.0), 1 mM EDTA, and 12.5 mM magnesium acetate. The mixture was denatured by heating to 95°C and annealed by cooling to 4°C in a thermal cycler (BioRad, Hercules, CA, USA). The fluorescence signal output of the encoding gate required the quencher strand to completely quench the fluorescent strand. The DNA sensing gate was visualized by 4% agarose gel electrophoresis in Tris-borate EDTA buffer for 40 min at 100 V and room temperature. The gel was visualized with a fluorescence scanner (E-Graph; ATTA, USA) (Fig. 15). Bands of the correct size (48 bp) of sensing gate were excised from the gel and purified using the Wizard DNA Clean-up System (Promega, Madison, WI, USA) according to the manufacturer's instructions.

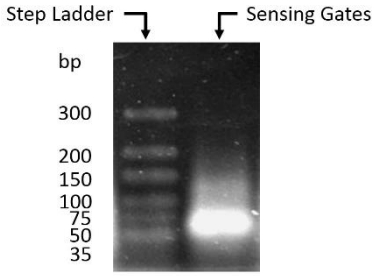



**Figure 15.** DNA analysis by gel electrophoresis.

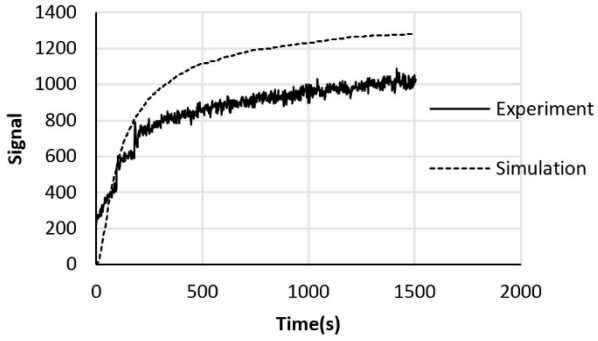



**Figure 16.** Hsa-miR-21 fluorophore signal intensity with respect to simulation results.

The output signal was detected using a Photonic Multichannel Analyzer spectrophotometer (PMA-12 C10027-02; Hamamatsu, Hamamatsu City, Japan) integrated with an Olympus IX71 microscope with a mercury UV lamp (Tokyo, Japan). Excitation and emission wavelengths of 580 and 608 nm, respectively, were used to detect ROX, and signals were captured over 1500 s. Hsa-miR-21 in the gates was detected as an increase in fluorescence signal intensity relative to the simulation results (Fig. 16). The signal output reflected rapid detection

and classification of the input. Moreover, when an input other than hsa-miR-21 was tested with the sensing and encoding system, the reactions produced no output signal.

## VI. CONCLUSION

A parallel decision-making system for a molecular robot was proposed, which showed that simple DNA strands could be used for decision-making from rule-based classifiers that classify a large input from binary tree classification by translating rule sets to DNA gates. The DNA strands are similar to chromosomes that increase the local concentration of DNA molecules, and can therefore be integrated into the molecular robot. The simulation results validated a set of rule-based classifiers for DNA computing. Simple experiments using hsa-miR-21 as input showed that the proposed system was effective in sensing and in signal output for a cancer class. Moreover, the parallel sensing and classification were rapid, sensitive, and selective for the input strand. In the future, the proposed system will be integrated into the molecular robot and used for decision making, and will also be tested in experiments using various miRNAs as input and as in vivo reactions in cancer cells.

## ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research (no. 22220001) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. The first author acknowledges financial support from the Directorate General of Higher Education, Ministry of Education and Culture of the Republic of Indonesia.

## REFERENCES

[1] C. Kaparissides, S. Alexandridou, K. Kotti, and S. Chaitidou, “Recent advances in novel drug delivery systems”. Journal of Nanotechnology, vol. 2, pp. 1–11, March 2006.

[2] M. Kumar, T. Ahmad, A. Sharma, U. Mabalirajan, A. Kulshreshtha, A. Agrawal, and G. Ghosh, “Let-7 microRNA-mediated regulation of IL-13 and allergic airway inflammation”. Journal of Allergy and Clinical Immunology, vol. 128, no. 5, pp. 1077–1085, November 2011.

[3] T. X. Lu, A. Munitz, and M. E. Rothenberg, "MicroRNA-21 is up-regulated in allergic airway inflammation and regulates IL-12p35 expression". The Journal of Immunology, vol. 182, no. 8, pp. 4994–5002, April 2009.

[4] T. X. Lu and M. E. Rothenberg, "Diagnostic, functional, and therapeutic roles of micro RNA in allergic diseases". Journal of Allergy and Clinical Immunology, vol. 132, no. 1, pp. 3–13, July 2013.

[5] X. Chen, Y. Ba, L. Ma, X. Cai, Y. Yin, K. Wang, J. Guo, Y. Zhang, J. Chen, X. Guo, Q. Li, X. Li, W. Wang, Y. Zhang, J. Wang, X. Jiang, Y. Xiang, C. Xu, P. Zheng, J. Zhang, R. Li, H. Zhang, X. Shang, T. Gong, G. Ning, J. Wang, K. Zen, J. Zhang, and C. Y. Zhang, "Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases". Cell Research, vol. 18, no. 10, pp. 997–1006, October 2008.

[6] C. Mavroidis and A. Ferreira, Nanorobotics: Current Approaches and Techniques, C. Mavroidis and A. Ferreira, eds. Springer, New York, NY, pp. 3, 2013.

[7] A. Ummat, A. Dubey, and C. Mavroidis, "Bionanorobotics: a field inspired by nature," in Y. Bar-Cohen, ed., Biomimetics: Biologically Inspired Technologies, CRC Press, Boca Raton, FL, pp. 201–227, 2005.

[8] K. Sanderson, "Bioengineering: What to make with DNA origami". Nature, vol. 464, no. 7286, pp. 158–159, March 2010.

[9] S. Hiyama, Y. Isogawa, T. Suda, Y. Moritani, and K. Sutoh. A design of an autonomous molecule loading/transporting/unloading system using DNA hybridization and biomolecular linear motors. " in Proc. European Nano Systems'05, pp. 75-80, 2005

[10] M. Hagiya, A. Konagaya, S. Kobayashi, H. Saito, and S. Murata, "Molecular Robots with Sensors and Intelligence". Accounts of Chemical Research, vol. 47, no. 6, pp. 1681–1690, 2014.

[11] G. Paun, G. Rozenberg, and A. Salomaa, DNA computing: new computing paradigms. G. Paun, G. Rozenberg, A. Salomaa, eds. Springer, New York, NY,pp: 40-65, 1998.

[12] M. N. Win and C. D. Smolke, "Higher-order cellular information processing with synthetic RNA devices". Science, vol. 322, no. 5900, pp. 456–460, October 2008.

[13] M. Hagiya, S. Wang, I. Kawamata, S. Murata, T. Isokawa, F. Peper, and K. Imai, "On DNA-Based Gellular Automata", in O. H. Ibarra, L. Kari, and S. Kopecki, eds., Unconventional Computation and Natural Computation, Springer International Publishing, New York, NY, pp. 177–189, 2014.

[14] S. Ayukawa, M. Takinoue, and D. Kiga, "RTRACS: a modularized RNA-dependent RNA transcription system with high programmability". Accounts of Chemical Research, vol. 44, no. 12, pp. 1369–1379, October 2011.

[15] D. Y. Zhang and G. Seelig, "Dynamic DNA nanotechnology using strand-displacement reactions". Nature Chemistry, vol. 3, no. 2, pp. 103–113, January 2011.

[16] G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree, "Enzyme-free nucleic acid logic circuits". Science, vol. 314, no. 5805, pp. 1585–1588, December 2006.

[17] D. Y. Zhang, A. J. Turberfield, B. Yurke, and E. Winfree, "Engineering entropy-driven reactions and networks catalyzed by DNA". Science, vol. 318, no. 5853, pp. 1121–1125, November 2007.

[18] L. Qian, E. Winfree, and J. Bruck, "Neural network computation with DNA strand displacement cascades". Nature, vol. 475, no. 7356, pp. 368–372, July 2011.

[19] L. Qian and E. Winfree, "Scaling up digital circuit computation with DNA strand displacement cascades". Science, vol. 332, no. 6034, pp. 1196–1201, June 2011.

[20] P. N. Tan, M. Steinbach, and V. Kumar. Introduction to data mining. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, Chapter 5.1, pp: 207 – 219, 2005.

[21] N. Rosenfeld, R. Aharonov, E. Meiri, S. Rosenwald, Y. Spector, M. Zepeniuk, H. Benjamin, N. Shabes, S. Tabak, A. Levy, D. Lebanony, Y. Goren, E. Silberschein, N. Targan, A. Ben-Ari, S. Gilad, N. Sion-Vardy, A. Tobar, M. Feinmesser, O. Kharenko, O. Nativ, D. Nass, M. Perelman, A. Yosepovich, B. Shalmon, S. Polak-Charcon, E. Fridman, A. Avniel, I. Bentwich, Z. Bentwich, D. Cohen, A. Chajut, I. Barshack, "MicroRNAs accurately identify cancer tissue origin". Nature Biotechnology, vol. 26, no. 4, pp. 462–469, April 2008.

[22] N. Kosaka, H. Iguchi, and T. Ochiya, "Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis". Cancer Science, vol. 101, no. 10, pp. 2087–2092, October 2010.

[23] D. Y. Zhang, "Dynamic DNA strand displacement circuits". Ph.D. dissertation, California Institute of Technology, 2010.

[24] P. W. Rothemund, "Folding DNA to create nanoscale shapes and patterns". Nature, vol. 440, no. 7082, pp. 297–302, January 2006.

[25] H. T. Maune, S. P. Han, R. D. Barish, M. Bockrath, W. A. Goddard III, P. W. Rothemund, and E. Winfree, "Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates". Nature Nanotechnology, vol. 5, no. 1, pp. 61–66, November 2009.

[26] S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, "Self-assembly of DNA into nanoscale three-dimensional shapes". Nature, vol. 459, no. 7245, pp. 414–418, May 2009.

[27] N. C. Seeman, "An overview of structural DNA nanotechnology". Molecular Biotechnology, vol. 37, no. 3, pp. 246–257, November 2007.

[28] S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih, "Rapid prototyping of 3D DNA-origami shapes with caDNAno". Nucleic Acids Research, vol. 37, no. 15, pp. 5001–5006, August 2009.

[29] CanDo - Computer-aided engineering for DNA origami, http://cando-dna-origami.org/, final request 19.11.2014.

[30] J. Liu, Z. Cao, and Y. Lu, "Functional nucleic acid sensors". Chemical Reviews, vol. 109, no. 5, pp. 1948–1998, May 2009.

[31] A. P. De Silva and S. Uchiyama, "Molecular logic and computing," Nature Nanotechnology, vol. 2, no. 7, pp. 399–410, 2007.

[32] Y. Benenson, "Biomolecular computing systems: principles, progress and potential". Nature Reviews Genetics, vol. 13, no. 7, pp. 455–468, July 2012.

[33] J. S. Shin and N. A. Pierce, "A synthetic DNA walker for molecular transport". Journal of the American Chemical Society, vol. 126, no. 35, pp. 10834–10835, September 2004.

[34] H. Qiu, J. C. Dewan, and N. C. Seeman, "A DNA decamer with a sticky end: the crystal structure of d-CGACGATCGT". Journal of Molecular Biology, vol. 267, no. 4, pp. 881–898, April 1997.

[35] S.H. Cha and C. Tappert, "A genetic algorithm for constructing compact binary decision trees". Journal of Pattern Recognition Research, vol. 4, no. 1, pp. 1–13, 2009.

[36] J. Han and M. Kamber. Data Mining, Southeast Asia Edition: Concepts and Techniques. Morgan Kaufmann, Burlington, MA,pp: 355-363, 2006.

[37] R. Garnier and J. Taylor, Discrete mathematics: proofs, structures and applications. CRC Press, Boca Raton, FL, 2009.

[38] Y. K. Lin and K. S. Fu, "Automatic classification of cervical cells using a binary tree classifier". Pattern Recognition, vol. 16, no. 1, 69–80, 1983.

[39] B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, and J. L. Neumann, "A DNA-fuelled molecular machine made of DNA". Nature, vol. 406, no. 6796, pp. 605–608, August 2000.

[40] H. Chandran, N. Gopalkrishnan, A. Phillips, and J. Reif, "Localized hybridization circuits," in L. Cardelli and W. Shih, eds., DNA Computing and Molecular Programming, Springer, Berlin, pp. 64–83, 2011.

[41] M. R. Lakin, S. Youssef, F. Polo, S. Emmott, and A. Phillips, "Visual DSD: a design and analysis tool for DNA strand displacement systems". Bioinformatics vol. 27, no. 22, pp. 3211–3213, November 2011.

[42] P. P. Medina, M. Nolde, and F. J. Slack, "Oncomir addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma". Nature, vol. 467, no. 7311, pp. 86–90, September 2010.

[43] NUPACK – Nucleic Acid Package. Available: http://www.nupack.org/, final request 19.11.2014.