nature portfolio

https://doi.org/10.1038/s42256-022-00502-7

Supplementary information

Molecular convolutional neural networks with DNA regulatory circuits

In the format provided by the authors and unedited

Molecular convolutional neural networks with DNA regulatory circuits

Supplementary information

Xiewei Xiong[1,3], Tong Zhu[1,3], Yun Zhu[1], Mengyao Cao[1], Jin Xiao[1], Li Li[1], Fei Wang[2], Chunhai Fan[2] and Hao Pei[1]*

[1]Shanghai Key Laboratory of Green Chemistry and Chemical Processes, School of Chemistry and Molecular Engineering, East China Normal University 500 Dongchuan Road, Shanghai 200241, China

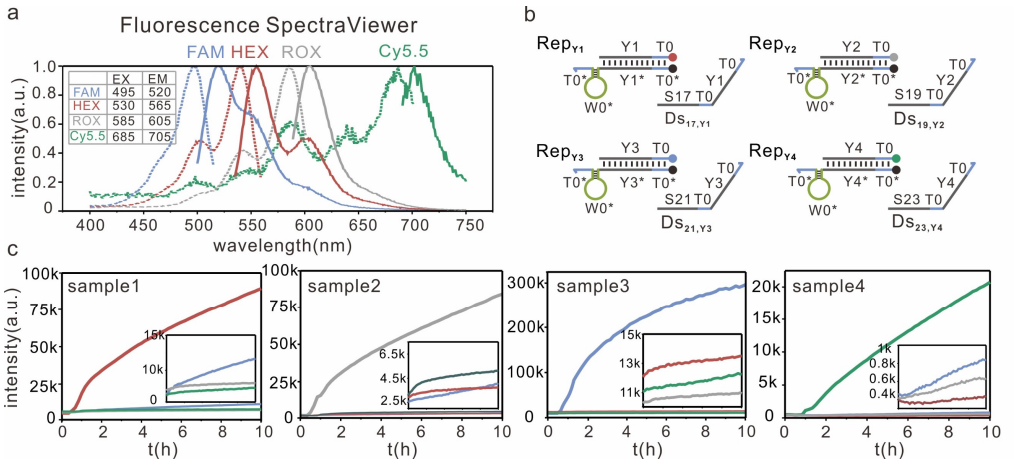
[2]School of Chemistry and Chemical Engineering, and Institute of Molecular Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University Shanghai 200240, China

[3]These authors contributed equally: Xiewei Xiong, Tong Zhu

*Corresponding author: peihao@chem.ecnu.edu.cn

Contents

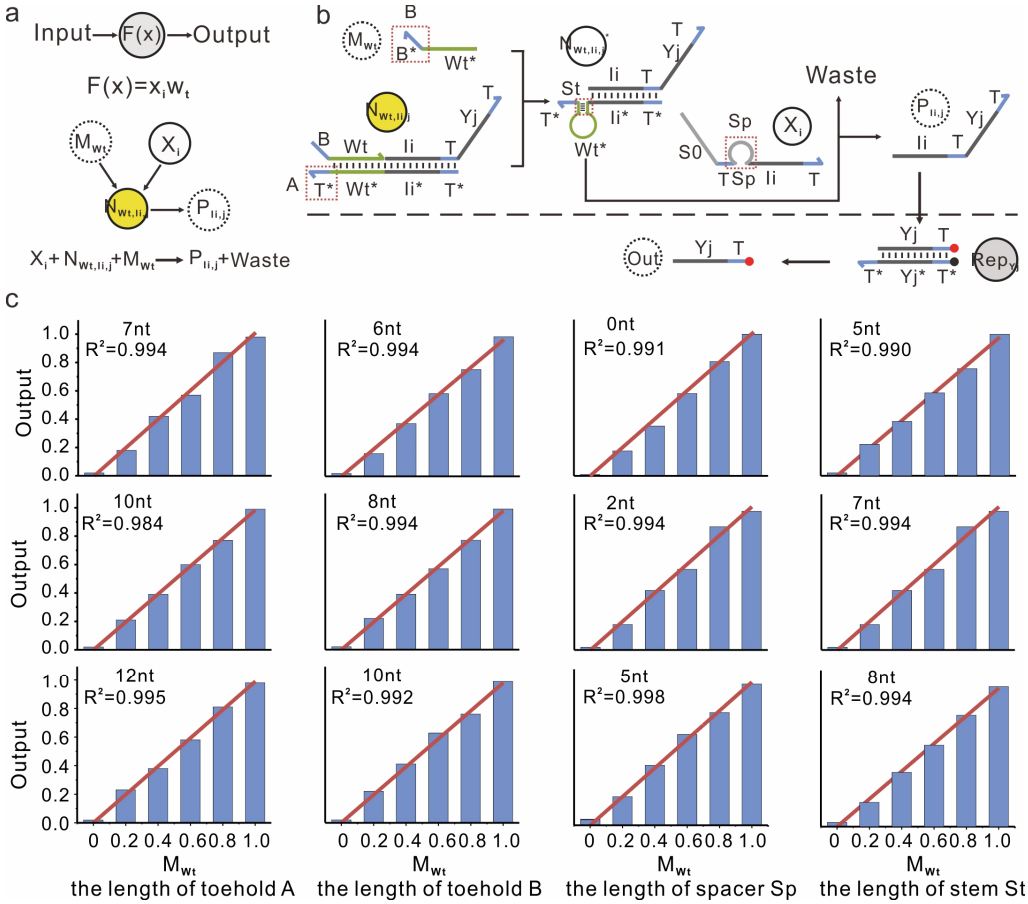
Reporter Crosstalk test. We have tested the crosstalk to verify the lack of significant interference between four different fluorophores before the fluorescence kinetics experiment (*Supplementary Fig. 1*). We mixed four reporters together, and prepared four solutions. Then, we added corresponding trigger strands to activate one reporter only. As displayed in *Supplementary Fig. 1c,* each channel only responded to the signal strand that triggers the corresponding reporter. Little crosstalk was observed in 10 h, with only ~6% crosstalk between the channels of FAM and HEX, which was still neglectable.



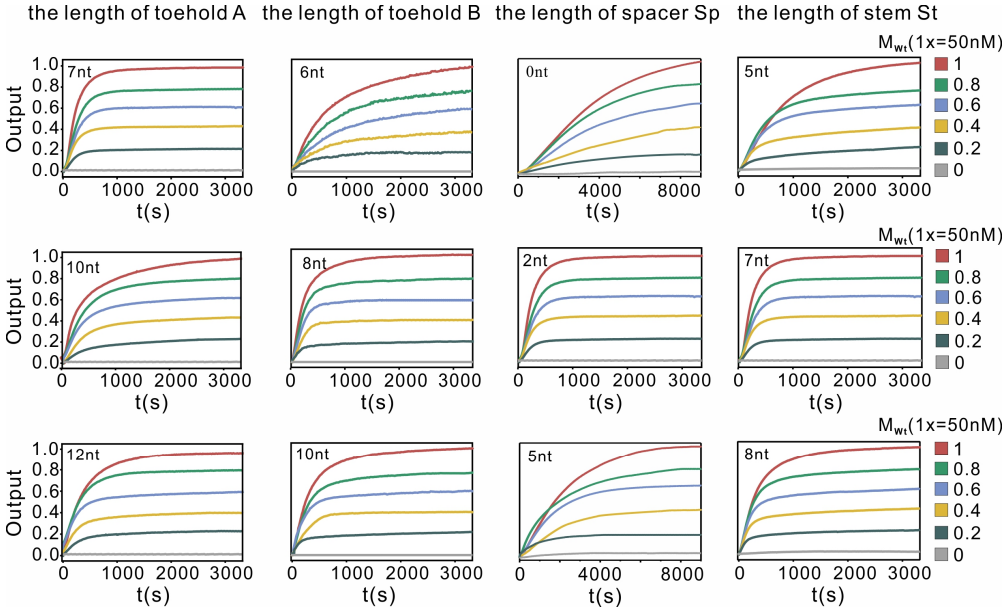
Supplementary Fig. 1| Crosstalk test among four reporters with distinct fluorophores. a, Fluorescence spectra (dotted curves: excitation spectra, solid curves: emission spectra) of FAM, HEX, ROX and Cy5.5. The inset table lists four sets of excitation and emission wavelengths that were determined to minimize the interference among different fluorophores. **b**, Four reporters with fluorophores FAM, HEX, ROX and Cy5.5, and four signal strands that can trigger the four reporters, respectively. **c**, Four parallel kinetics experiments to test the crosstalk among the four reporters. Four channels were set with the wavelengths listed in the inset table in *Supplementary Fig. 1a*. All four reporters were added to each of four cuvettes at the beginning of the experiment (t=0). Four signal strands were then added to four cuvettes separately (roughly t=25 mins). Matching colors were used to visualize corresponding trajectories and outputs.

S1. Robustness of weight multiplications

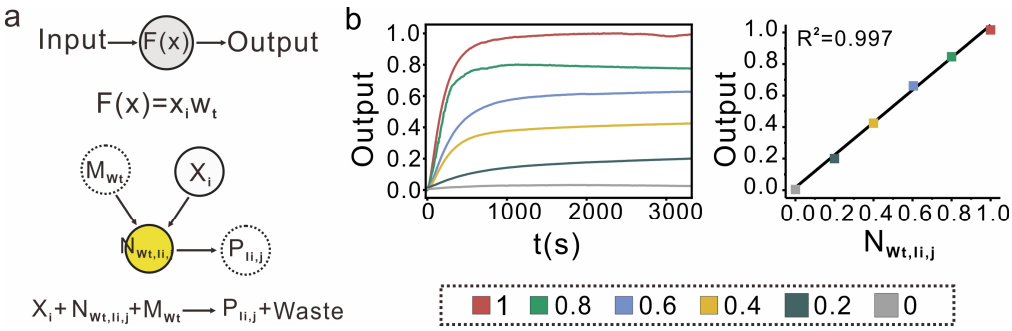
We have tested the DNA implementation of weight multiplication under different conditions, including the length of toehold domains (toehold A and toehold B), the length of stem domains (stem domain St), and the length of spacer domains (spacer domain Sp) (*Supplementary Figs. 2 and 3*). As shown in *Supplementary Fig. 2c*, we observed the linear relationship between the output signal of weighted multiplication and concentration of M_{Wt} ($R^2 > 0.98$) for various domain lengths. Note that we used 15-nucleotide loop domain in initial experiments (*Supplementary Figs. 2-4*). We also verified the DNA implementation of weight multiplication with 13-nucleotide loop domain design (*Supplementary Fig. 5*). For the simplification of sequence design, we used 13-nucleotide loop domain for the rest experiments. Furthermore, we observed the linear relationship between the concentration of weight tuning molecule M_{Wt} and the output signal under different temperatures (25 ℃ and 20 ℃) (*Supplementary Fig. 6*) and molecular concentrations (*Supplementary Fig. 7*). These results lead to three main observations: First, switching function——weight substrate molecule $N_{Wt,Ii,j}$ is OFF unless the addition of weight tuning molecule M_{Wt}. Second, the value of weights can be facilely regulated by the concentration of M_{Wt} without changing the components. Third, the sequence of weight tuning domains can be independently designed and only associated with the assignment of weights.



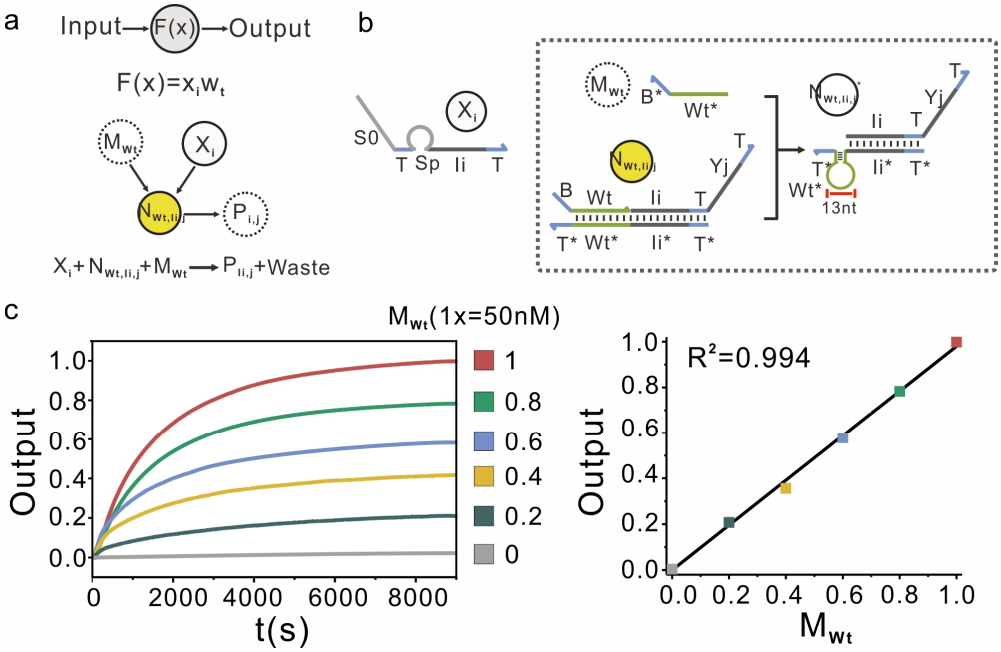
Supplementary Fig. 2| The DNA implementation of weight multiplication under different conditions. a, The abstract schematic of the weight multiplication. **b**, The DNA implementation of weight multiplication, with 15-nucleotide loop domain. **c**, Kinetics experiments of different length of domains on the operation of the weight multiplication. The concentration of the weight substrate molecule $N_{Wt,Ii,j}$ and input X_i are 2×, and the concentration of the reporter Rep_{Yj} is 2×. The standard concentration is 50 nM (1× = 50 nM).



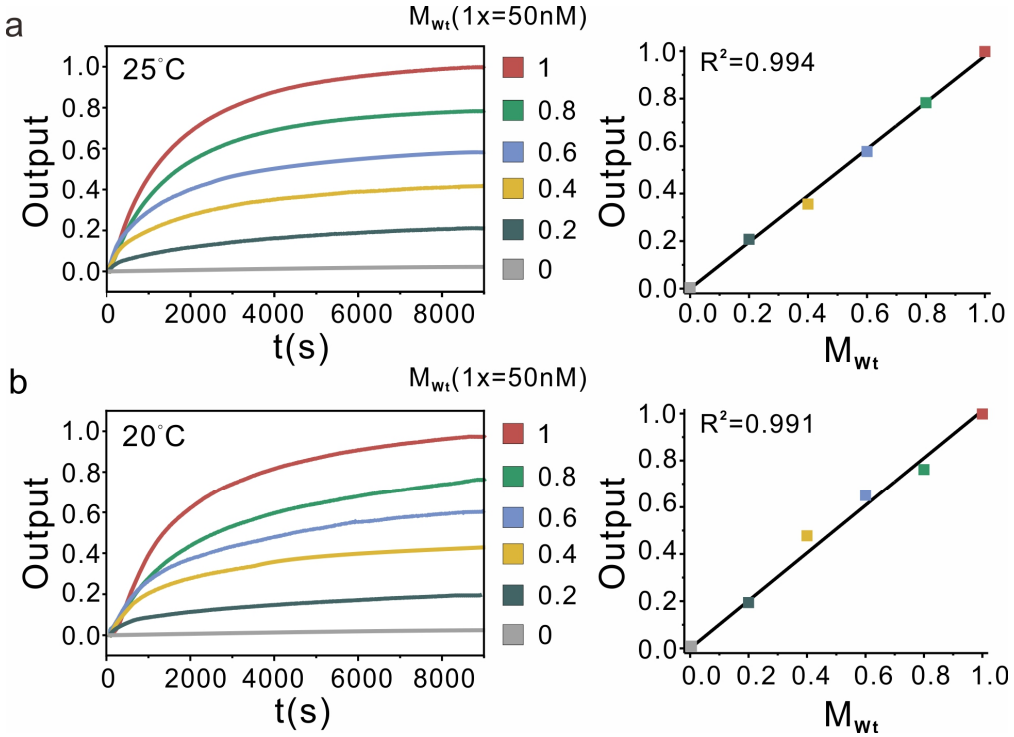
Supplementary Fig. 3| Fluorescence kinetics data of the weight multiplication shown in *Supplementary Fig. 2*. All fluorescence kinetics data are shown over the course of 3600s.



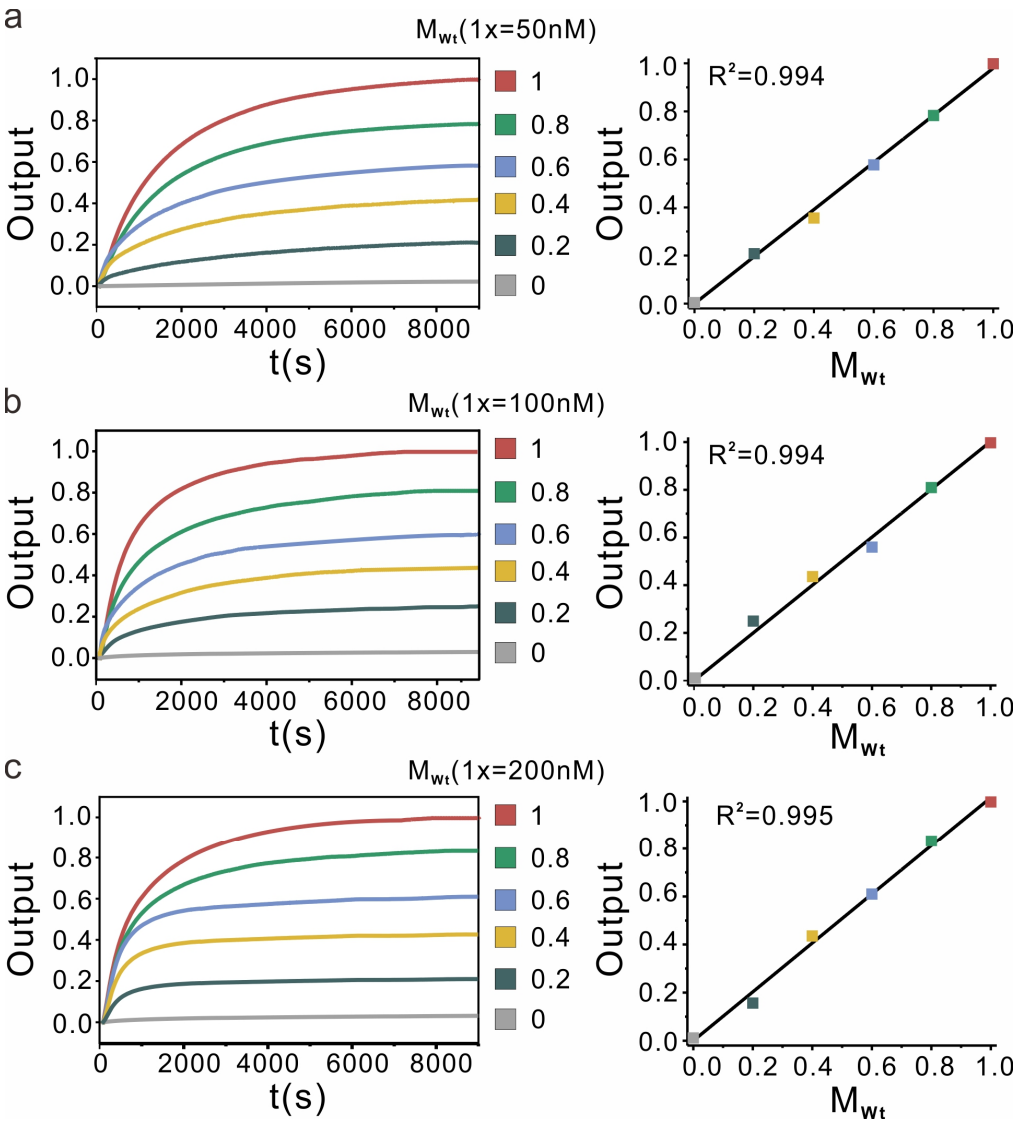
Supplementary Fig. 4| a, The abstract schematic of the weight multiplication. **c**, Kinetics experiments of the weight multiplication. The concentration of the weight tuning molecule M_{Wt} and input X_i are 2×, and the concentration of the reporter Rep_{Yj} is 2×. The standard concentration is 50 nM (1× = 50 nM).



Supplementary Fig. 5| The DNA implementation of weight multiplication. a, The abstract schematic of the weight multiplication. **b**, The DNA implementation of weight multiplication, with 13-nucleotide loop domain and 6-nucleotide toehold domain. **c**, Fluorescence kinetics data (left) and the steady state fluorescence response (right) of the weight multiplication. The concentration of weight substrate molecule $N_{Wt,Ii,j}$ and input X_i are 2×, and the concentration of the reporter Rep_{Yj} is 2×. The standard concentration is 50 nM (1× = 50 nM).



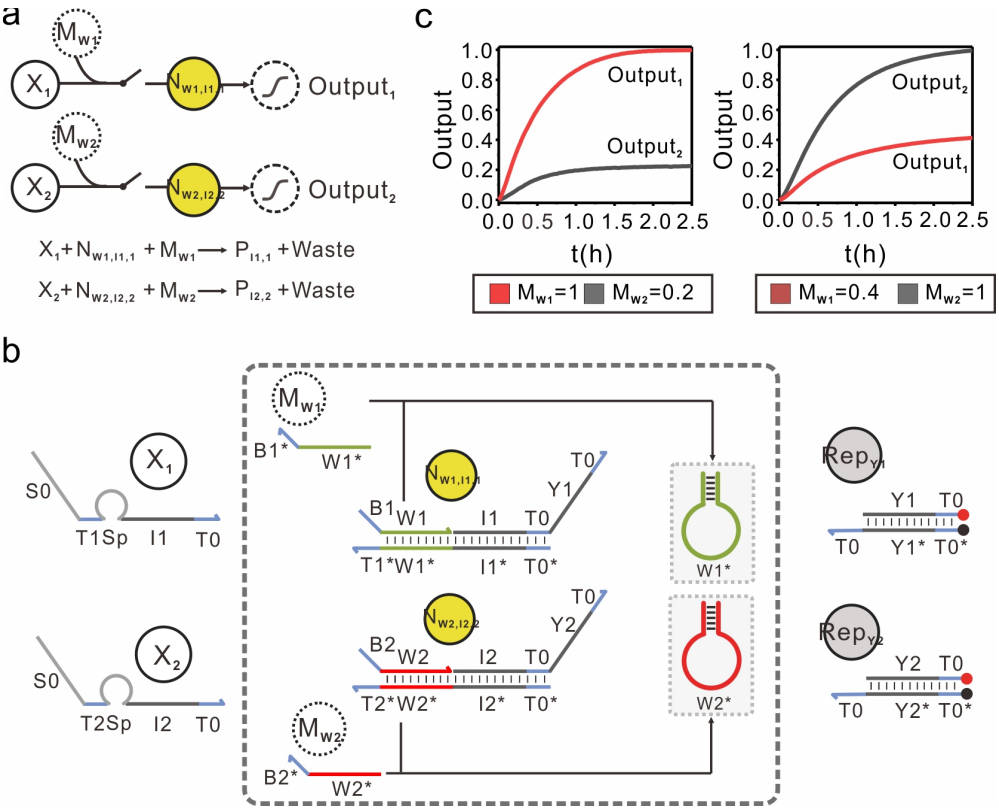
Supplementary Fig. 6| Kinetics experiments of reaction temperature (a) 25 ℃ and (b) 20 ℃ on the operation of the weight multiplication shown in *Supplementary Fig. 5a*. Fluorescence kinetics data (left) and the steady state fluorescence response (right) of the weight multiplication. The concentration of weight substrate molecule $N_{Wt,Ii,j}$ and input X_i are 2×, and the concentration of the reporter Rep_{Yj} is 2×. The standard concentration is 50 nM (1× = 50 nM).



Supplementary Fig. 7| Kinetics experiments of different concentrations on the operation of the weight multiplication shown in *Supplementary Fig. 5a*. (a-c) Fluorescence kinetics data (left) and the steady state fluorescence response (right) of the weight multiplication. The concentration of weight substrate molecule $N_{Wi,Ii,j}$ and input X_i are 2×, and the concentration of the reporter Rep_{Yj} is 2×.

S2. Independent molecular weight multiplication systems

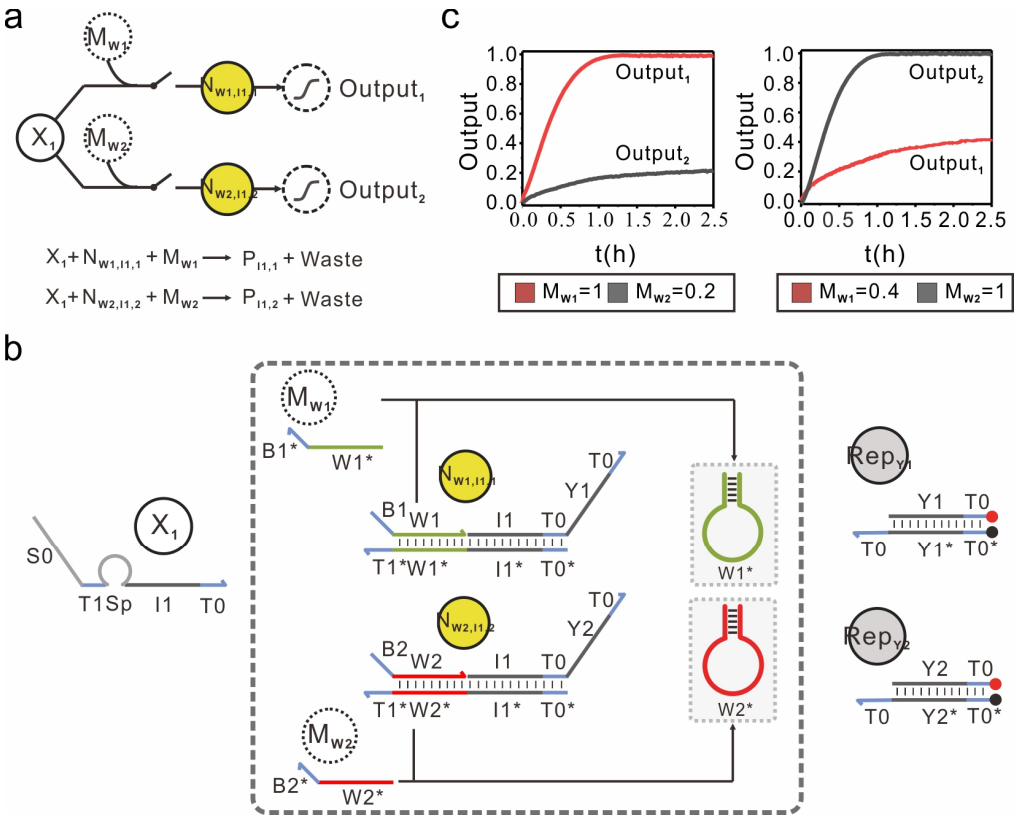
For scaling up to large circuits, independent molecular weight multiplication systems must have negligible crosstalk. Two independent weight multiplication systems (w_1x_1 and w_2x_2) were designed and tested in the same solution (*Supplementary Fig. 8a*). First, the two inputs are encoded with two different DNA input strands (X_1 and X_2 in *Supplementary Fig. 8b*); the weight is encoded in the sequence of weight tuning domains (see green domains W1* and red domains W2* in *Supplementary Fig. 8b*). To complete the two multiplication, we then designed two weight substrate molecules ($N_{W1,I1,Y1}$ and $N_{W2,I2,Y2}$ in *Supplementary Fig. 8b*), which have different weight tuning domains corresponding to two weights, and have different recognition domains (I1* and I2*) to respond to different inputs (X_1 and X_2). The values of two weights w_1 and w_2 are implemented by concentrations of two different weight tuning molecules M_{Wt}. Two distinct fluorescent reporters (FAM and ROX) were used to report corresponding outputs. As shown in *Supplementary Fig. 8c*, two fluorescence steady-state signals are linearly proportional to the concentration of M_{Wt} (M_{W1} and M_{W2}). These results suggest that each appliance functions modularly, not affecting or being affected by the function of others.



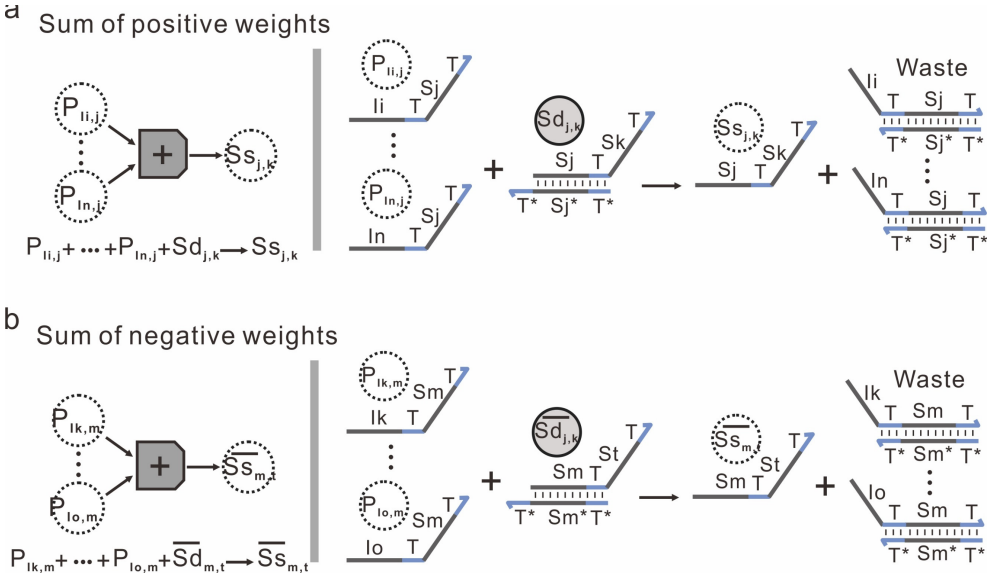
Supplementary Fig. 8| Two independent weight multiplication systems. a, The abstract schematic of two independent weight multiplication systems. **b**, The DNA implementation of two independent weight multiplication systems. **c**, Fluorescence kinetics data of two independent weight multiplication systems with different concentrations of M_{W1} and M_{W2}. Concentrations of weight substrate molecules $N_{Wt,Ii,j}$ and inputs X_i are 2×, and concentrations of reporters Rep_{Yj} is 2×. The standard concentration is 50 nM (1× = 50 nM).

Given that the sequence of weight tuning domains is independently designed and correlated with the assignment of weights, independent weight multiplication systems can thus be built with different weight tuning domain sequences design. As shown in *Supplementary Fig. 9a*, two independent weight multiplication systems that can interact with the same input X_1 were designed and tested in the same solution. The weights are encoded in the sequence of weight tuning domains (see green domains W1* and red domains W2* in *Supplementary Fig. 9b*). To complete the two multiplication, we then designed two weight substrate molecules ($N_{W1,I1,Y1}$ and $N_{W2,I1,Y2}$ in *Supplementary Fig. 8b*), which have different weight tuning domains corresponding to two weights, but have the same recognition domains (I1*) to respond to one input (X_1). The values of two weights w_1 and w_2 are determined by two different weight tuning molecules M_{Wt}. As shown in *Supplementary Fig. 9c*, two fluorescence steady-state signals were found to be linearly proportional to the concentration of M_{Wt} (M_{W1} and M_{W2}). Collectively, these results suggest that the

assignment of weights is only associated with the weight tuning domain sequences, wherein different assigned weights can be realized by simply changing sequences of the weight tuning domain.



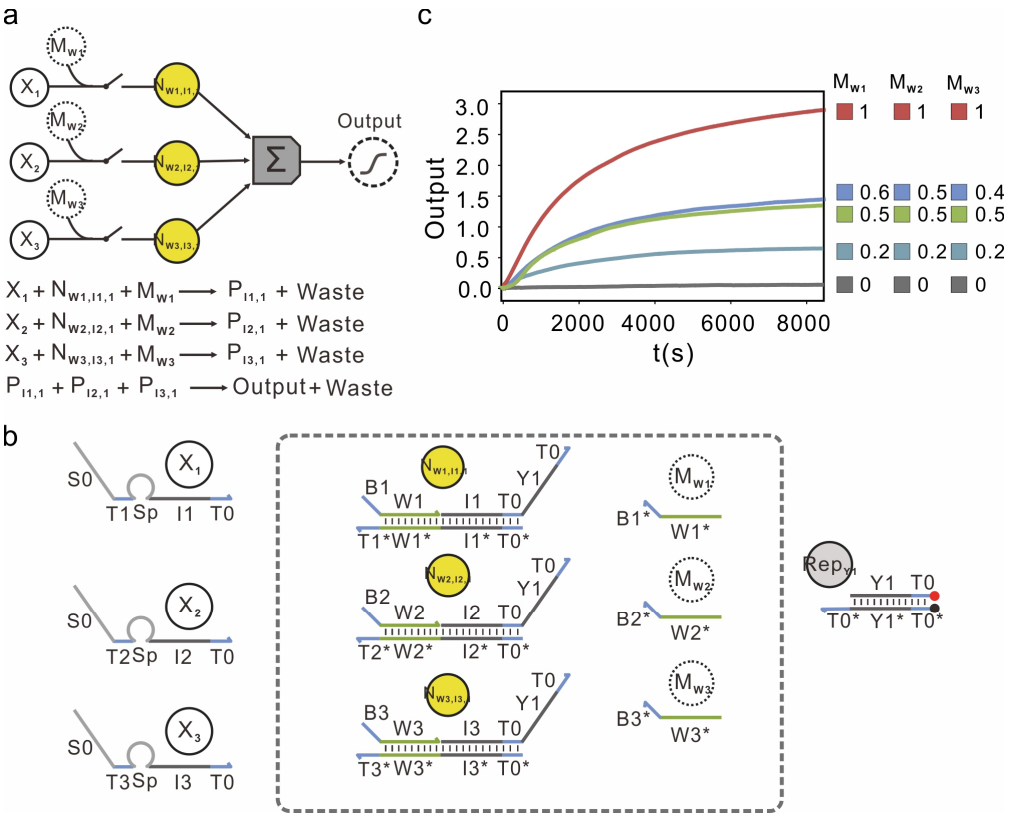
Supplementary Fig. 9| Two independent weight multiplication systems. a, The abstract schematic of two independent weight multiplication systems. **b**, The DNA implementation of two independent weight multiplication systems. **c**, Fluorescence kinetics data of two independent weight multiplication systems with different concentrations of weight tuning molecules M_{W1} and M_{W2}. Concentrations of weight substrate molecules $N_{Wt,Ii,j}$ and reporters Rep_{Yj} are $2\times$, and concentrations of input X_i is $4\times$. The standard concentration is 50 nM ($1\times = 50$ nM).



Supplementary Fig. 10| Summation can be implemented with a simple chemical reaction (listed in the left) and DNA strand displacement process (listed in the right). Different summation gates are designed to implement the summation, and weights with positive and negative values are implemented using different output sequence (for example, **a,** $P_{Ii,j}$ and **b,** $P_{Ik,m}$).

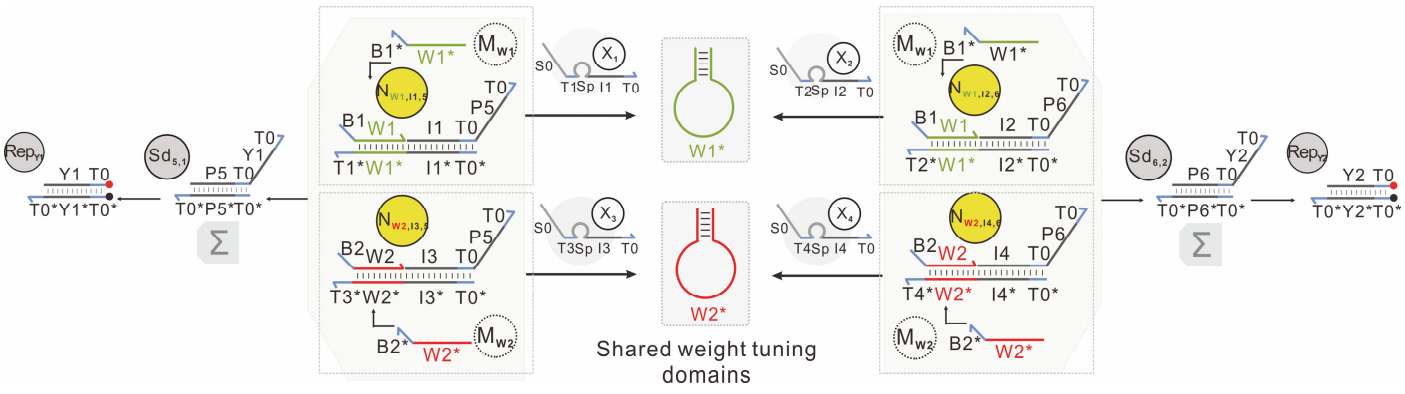
S3. The DNA implementation of multiply-accumulate operation

Three-input multiply-accumulate (MAC) operation $y = w_1x_1 + w_2x_2 + w_3x_3$ can be implemented at a molecular level (*Supplementary Fig. 11a*). Three different weight tuning molecules M_{Wt} with different sequences of weight tuning domains were used to set the value of weights, while the same output sequence of the intermediate species $P_{Ii,j}$ was designed to sum up weighted inputs (*Supplementary Fig. 11b*). As shown in *Supplementary Fig. 11c*, the output signal y was proportional to the weight assigned to each input, suggesting that the circuit was performed correctly. Taking these results together, we believe this summation operation can accurately determine the weighted sums from the same neuron.

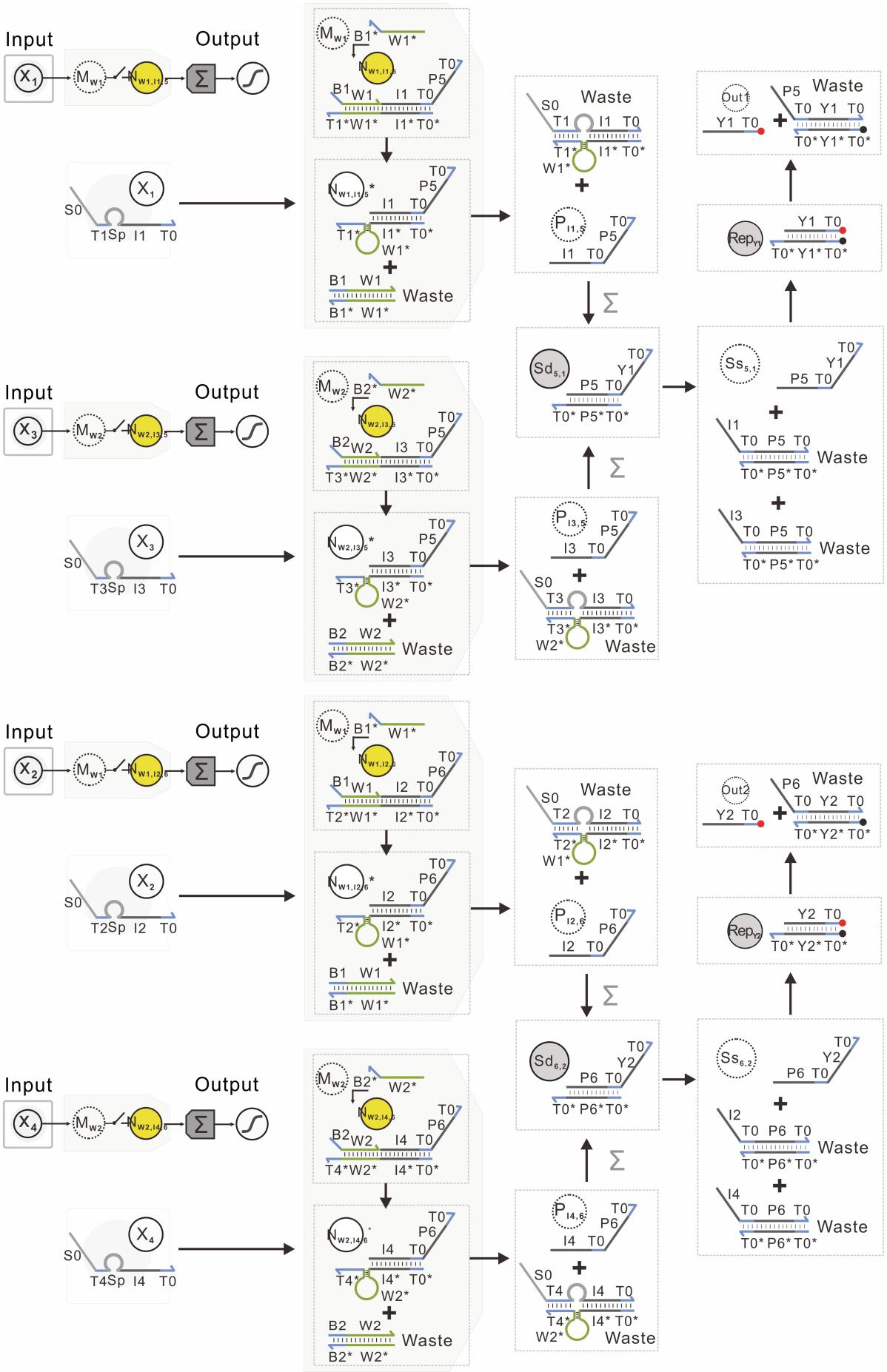


Supplementary Fig. 11| The DNA implementation of three-input MAC operation. a, The abstract schematic of three-input MAC operation. **b**, The DNA implementation of three-input MAC operation. **c**, Fluorescence kinetics data of three-input MAC operation with different concentrations of weight tuning molecules M_{W1}, M_{W2} and M_{W3}. Concentrations of weight substrate molecules $N_{Wt,Ii,j}$ and inputs X_i are 2×, and concentration of the reporter Rep_{Yj} is 6×. The standard concentration is 50 nM (1× = 50 nM).

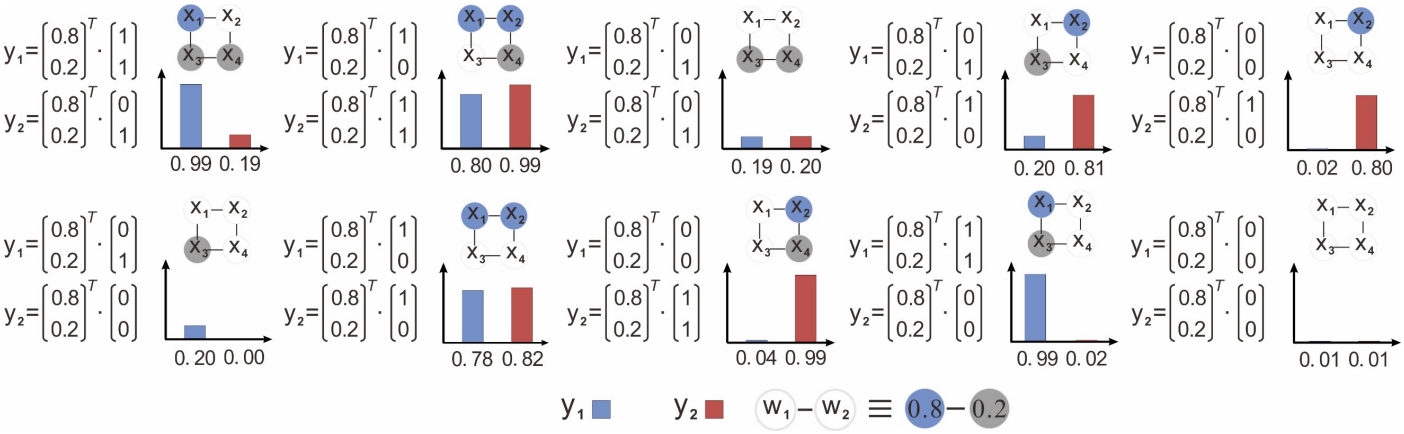
S4. The DNA implementation of convolution operation



Supplementary Fig. 12| The DNA implementation of convolution operation of a 2×2 input pattern using a 2×1 kernel. The DNA circuit consists of four weight substrate molecules $N_{Wt,Ii,j}$, two weight tuning molecules M_{Wt}, two summation gates $Sd_{j,k}$, and two reporter Rep_{Yi}. Four inputs are encoded in the sequence of recognition domain Ii, in which each 1 or 0 corresponded to the presence or absence of an input strand X_i. Two shared weights were encoded in the sequence of two weight tuning domains (green domain W1* and red domain W2*), in which two weight tuning molecules M_{W1} and M_{W2} can be used to activate four weight substrate molecules ($N_{W1,I1,5}$, $N_{W1,I2,6}$, $N_{W2,I3,5}$, and $N_{W2,I4,6}$). Note that $N_{W1,I1,5}$ and $N_{W1,I2,6}$ have the same weight tuning domain W1*, while $N_{W2,I3,5}$ and $N_{W2,I4,6}$ have the same weight tuning domain W2*.

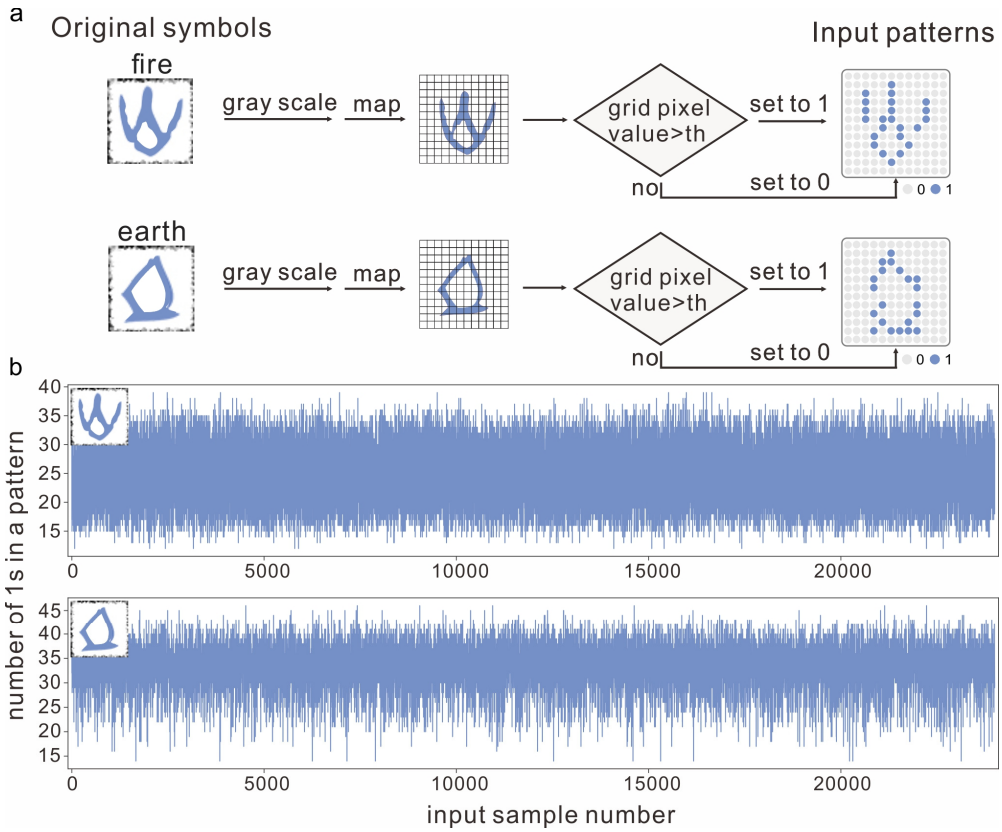


Supplementary Fig. 13| The detailed stepwise convolution operation of a 2×2 input pattern using a 2×1 kernel. Left: Abstract schematic diagram of a MAC operation with one input. Right: Stepwise illustration of the mechanism of components interaction from an Input to a reporter. Arrows indicate flows of the reactions.



Supplementary Fig. 14| Characterization of convolution with ten different input patterns after 3 h. The value of weights determined from convolution kernel is 0.8× (blue dot) and 0.2× (gray dot), respectively. The blue and red histograms correspond to the fluorescence kinetic outputs of y_1 and y_2, respectively. Blue dots indicate that the weight w_1 that is multiply with corresponding input (X_1 or X_2), while gray dots indicate that the weight w_2 that is multiply with corresponding input (X_3 or X_4).

S5. A DNA-based ConvNet for one of two rotated molecular patterns recognition

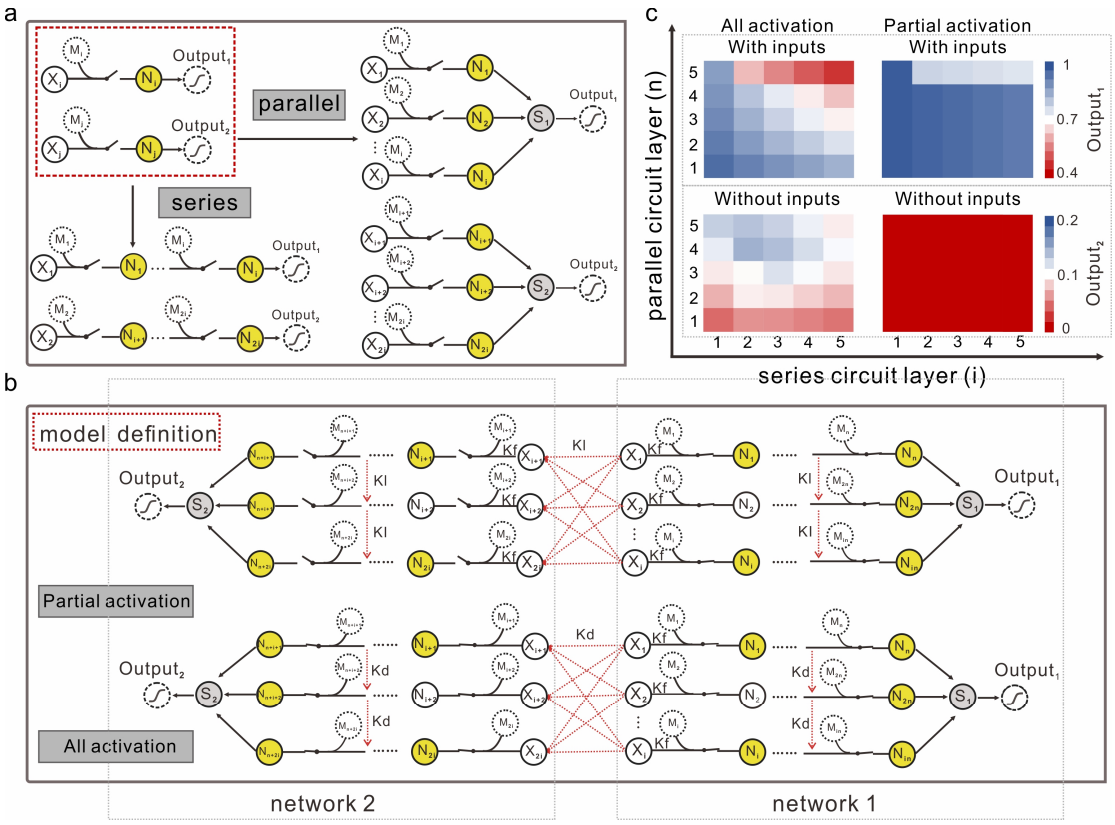


Supplementary Fig. 15| Determination of two input patterns. a, The symbols were rescaled to a 12×12 grid. The bits whose grid pixel value were larger than a threshold was set to 1, while the remaining bits were set to 0. Then binary input patterns of 'fire' and 'earth' were obtained. **b**, The number of 1s in each input pattern. The number of 1s in the 'fire' input pattern is mostly between 15 and 35, while the number of 1s in the 'earth' input pattern is mostly between 25 and 40.

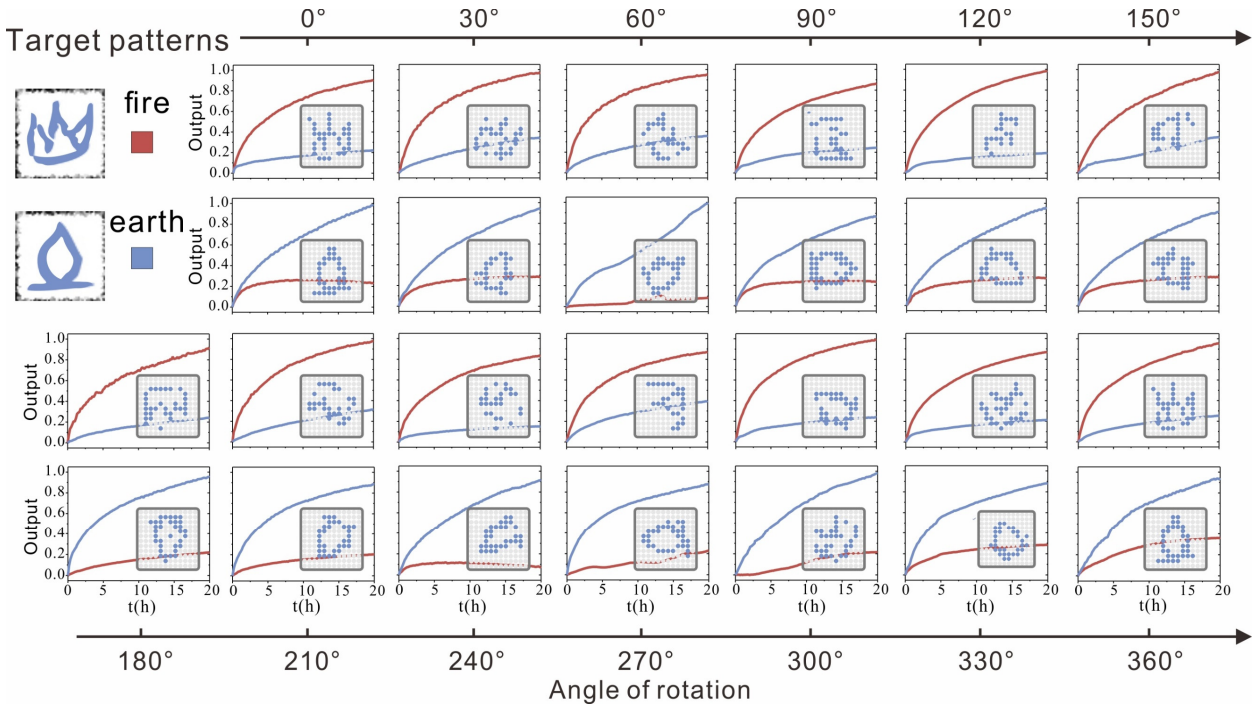
One desirable feature of the molecular weight multiplication is switching function, which in theory can improve the interference robustness of the DNA molecular network. This allows DNA circuits to be able to switch a specific reaction pathway in the convolution layer when exposed to a specific pattern. A quantitative model has been proposed to verify the interference robustness of DNA reaction network (*Supplementary Fig. 16*). It assumes crosstalk is unavoidable, that is, non-specific displacement may occur when two different domains hybridize to their respective complements in the DNA reaction network (the reaction rate constant for single strand and non-target complex is K_d, and set as 2×10^3).

In the quantitative model, we gradually scale up the network (network 1 and network 2) by cascading the weight multiplication in parallel and in series (*Supplementary Fig. 16a, b*). We investigated the signal loss of network 1 with all and partial activated reaction pathway in the presence of inputs; we also

investigated the signal leakage of network 2 with all and partial activated reaction pathway in the absence of inputs. As shown in *Supplementary Fig. 16b*, the reaction rate constant for single strand and non-target complex is K_d and K_l for networks that were all activated and partially activated, respectively. For partial activation, weight substrate molecule N_i is inactivated, resulting in only zero-toehold strand displacements with the non-specifically target domains. Two networks were simulated in Visual DSD[1], in which the embedded stochastic simulator yields the concentration trajectories of two outputs, using corresponding rate constants K_m, K_d and K_l. As shown in *Supplementary Fig. 16c*, we found that the signal loss of network 1 in the presence of inputs was only about 20% when the cascading layer is 5 (i=5, n=5) with partial activated reaction pathway. However, the signal loss of network 1 was up to 54% when the cascading layer is 5 (i=5, n=5) with all activated reaction pathway. In addition, we found that the signal leakage of network 2 in the absence of inputs was still about 0% when the cascading layer is 5 (i=5, n=5) with partial activated reaction pathway, while the signal leakage of network 2 was up to 15% when the cascading layer is i=2 and n=4 with all activated reaction pathway. These results revealed that activating a specific set of wires in the reaction network could yield a better anti-interference behavior.

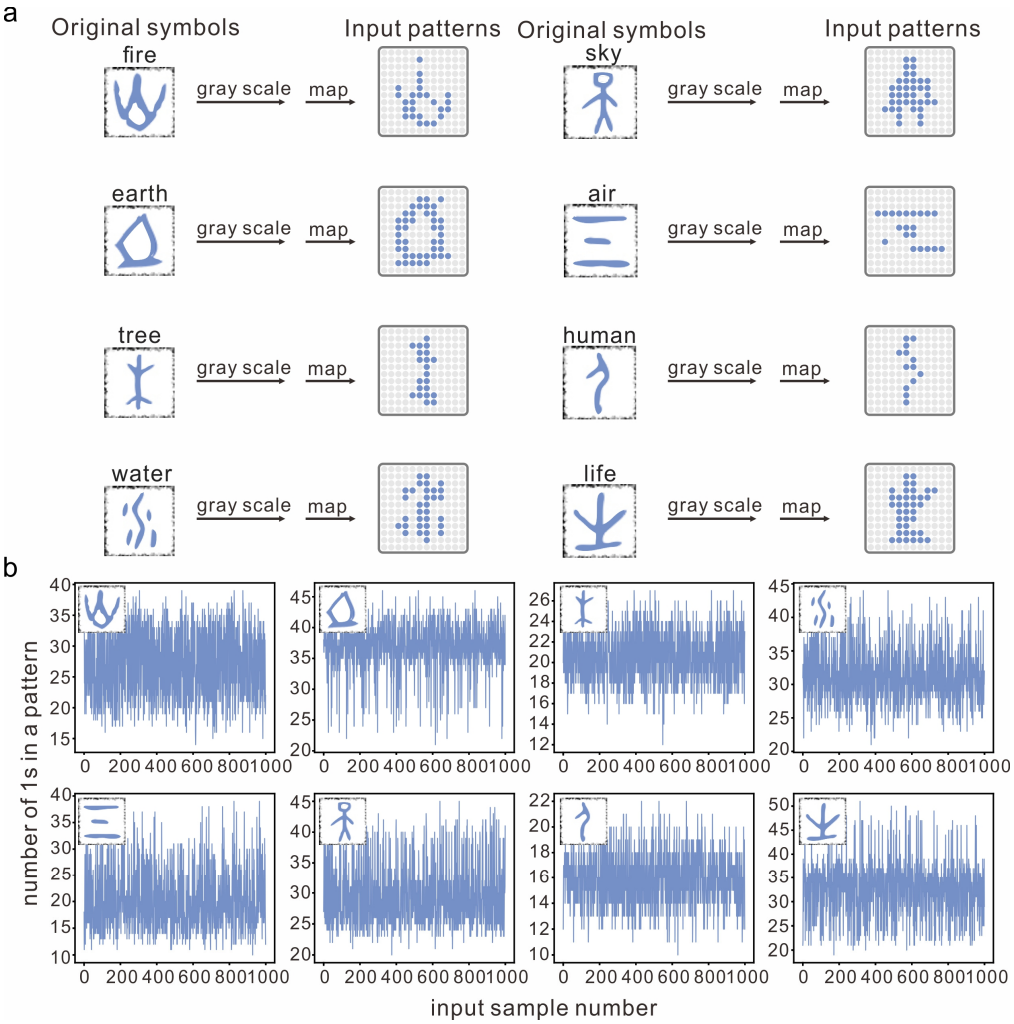


Supplementary Fig. 16| A quantitative model has been proposed to verify the interference robustness by Visual DSD. **a**, The network can be scaled up by cascading the weight multiplication in parallel and in series. **b**, The model definition. The desired reaction rate between the target domains K_m=1.5×10^5 M^{-1}s^{-1}, the reaction rate between non-target strands K_d = 2×10^3 M^{-1}s^{-1}, and the reaction rate of zero-toehold strand displacement K_l=1 M^{-1}s^{-1}. In panels **a** and **b**, the quantities M_{Wi} and $N_{Wi,Ii,j}$ have been shortened to M_i and N_i for the sake of brevity. **c**, The standard signals of two outputs with increasing circuit layers. Two networks were simulated in Visual DSD, the embedded stochastic simulator plots the concentration trajectories of two outputs, using corresponding rate constants K_m, K_d and K_l.

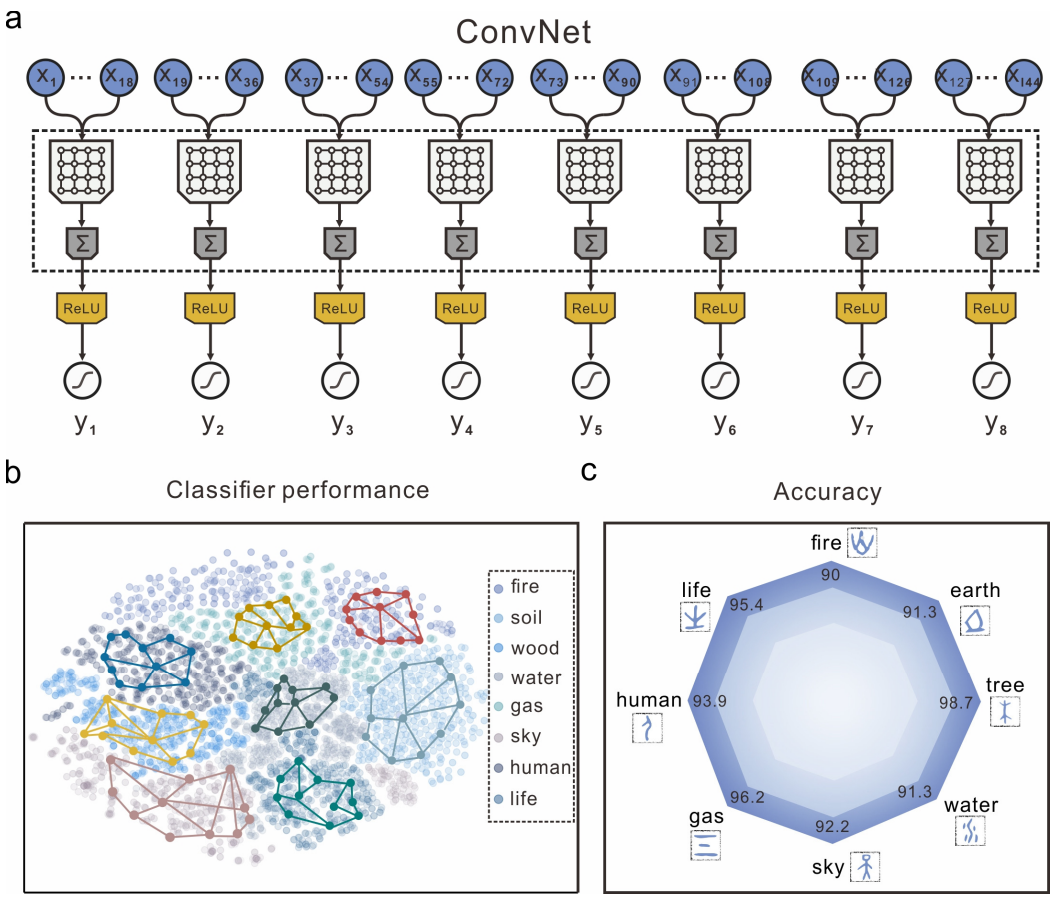


Supplementary Fig. 17| Fluorescence kinetic experiments to characterize the recognition behavior of 'fire' (red curves) and 'earth' (blue curves) with rotation angle from 0° to 360° with the DNA-based ConvNet.

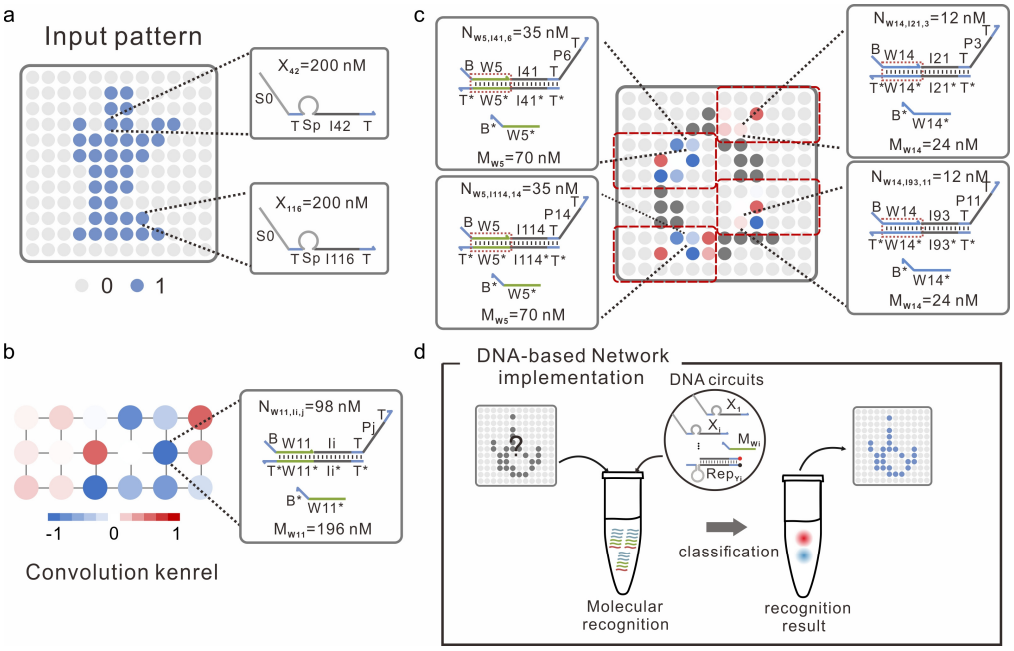
S6. A DNA-based ConvNet for one of eight molecular patterns recognition



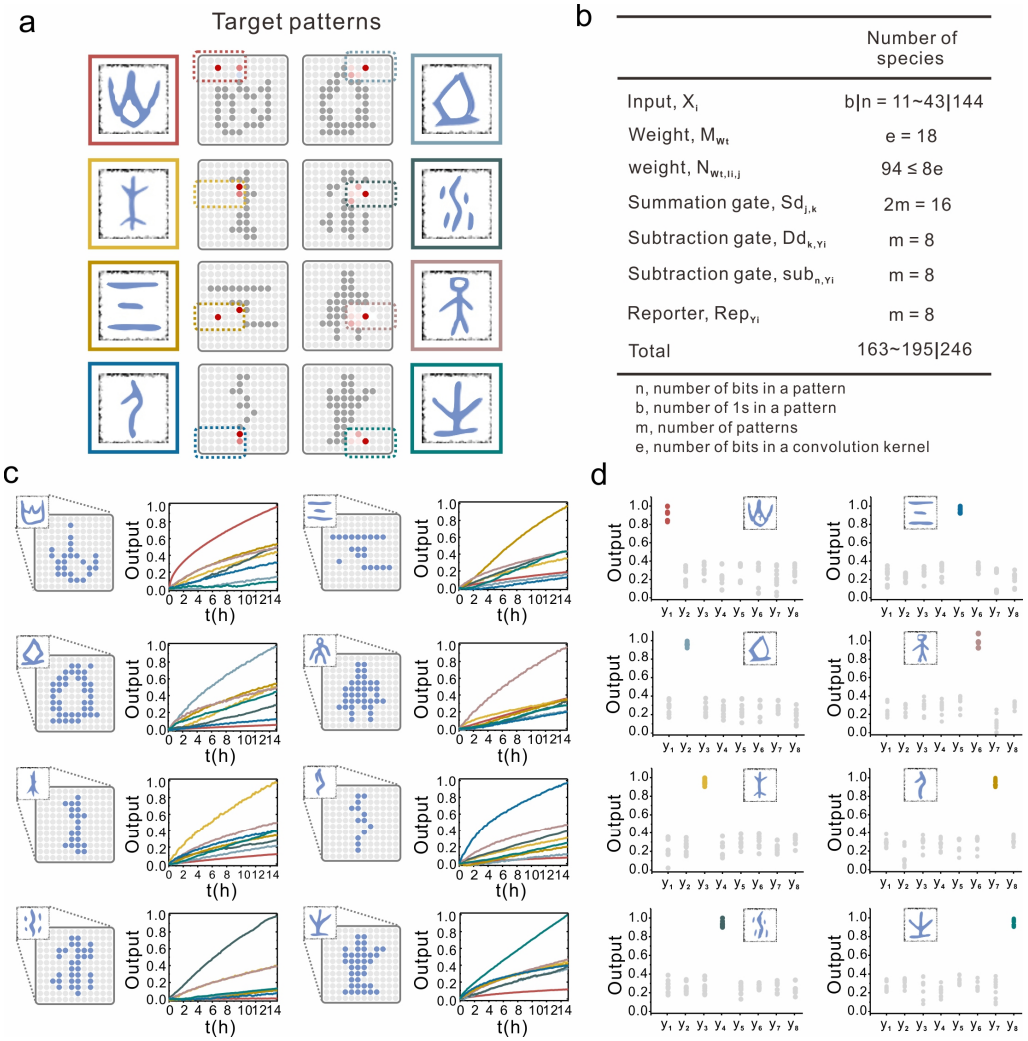
Supplementary Fig. 18| Determination of eight input patterns. a, The symbols were rescaled to a 12×12 grid. The bits whose grid pixel value were larger than a threshold were set to 1, while the remaining bits were set to 0. Then, the binary input patterns of eight oracles were obtained. **b**, The number of 1s in each input pattern. The number of 1s in the 'fire' input pattern is mostly between 20 and 35. The number of 1s in the 'earth' input pattern is mostly between 35 and 40. The number of 1s in the 'tree' input pattern is mostly between 18 and 23. The number of 1s in the 'water' input pattern is mostly between 28 and 35. The number of 1s in the 'sky' input pattern is mostly between 15 and 25. The number of 1s in the 'air' input pattern is mostly between 25 and 35. The number of 1s in the 'human' input pattern is mostly between 14 and 18. The number of 1s in the 'life' input pattern is mostly between 25 and 37.



Supplementary Fig. 19| The ConvNet for one of eight molecular patterns recognition. a, Circuit diagram for recognizing eight distinct patterns. **b**, Performance of the ConvNet in silico for the reference dataset with eight oracles. **c**, The radar map of the accuracy of the ConvNet. where 90.0%, 91.3%, 98.7%, 91.3%, 92.2%, 96.2%, 93.9%, and 95.4% of 'fire', 'earth', 'tree', 'water', 'sky', 'gas', 'human' and 'life' were recognized correctly in theory.



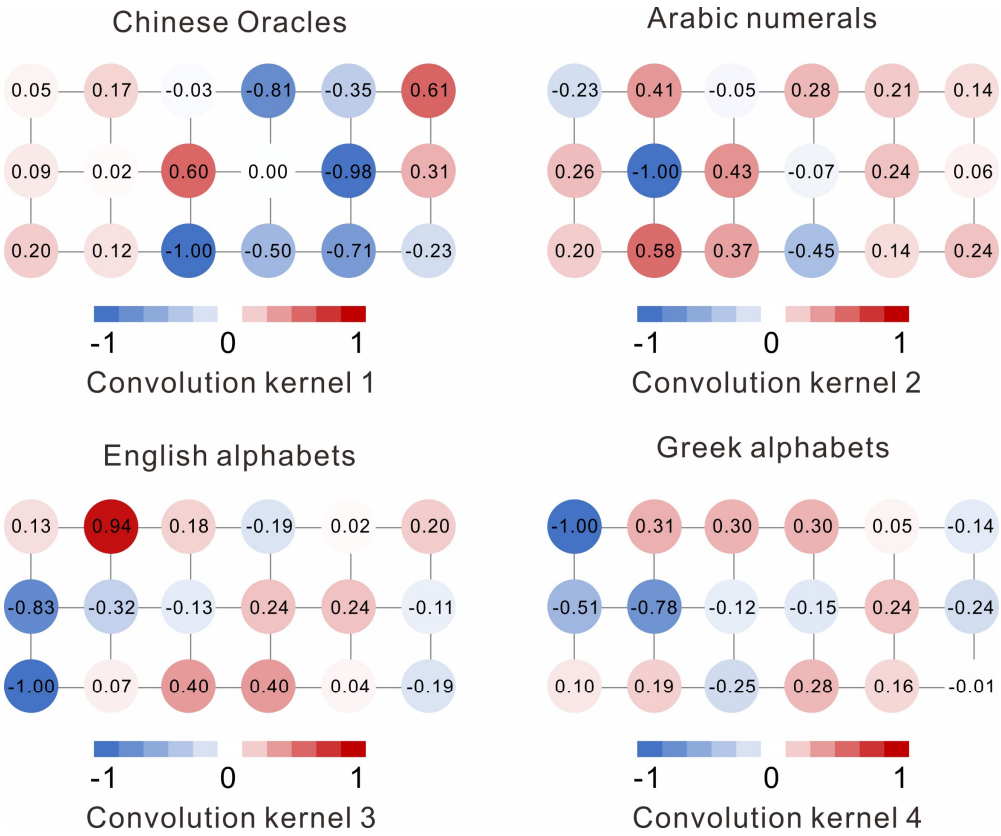
Supplementary Fig. 20| a, Example binary inputs with each 1 and 0 corresponding to the presence and absence of an input strand, respectively. The concentration of each input strand is 200 nM. **b**, The DNA implementation of the convolution kernel. The concentration of each weight substrate molecule (for example, 98 nM for the 11th pixel) was determined by the weight value of the convolution kernel. **c**, The shared convolution kernel reacts with each receptive region to implement the weight multiplication. The value of each pixel in each receptive region was used to determine concentrations of each weight substrate molecule. Because of the shared convolution kernel, the sequence of weight tuning domain (green region) of weight species is the same for each pixel that interacts with the same kernel function in different receptive regions. **d**, Unknow input pattern was added to the solution, upon addition of the DNA circuit, fluorescence signals were readout to report recognition results.



Supplementary Fig. 21| A DNA-based ConvNet for one of eight molecular patterns recognition. a, The characteristic receptive region of different target patterns (colors are associated with respective output trajectories in **c** and **d**). **b**, The number of distinct species in the circuit. The bottom row lists the number of species for a specific number of inputs (left) and for all possible inputs (right). We chose representative example patterns with up to 43 1s from the dataset, which corresponds to up to 43 distinct single strands. **c**, Fluorescence kinetics data of the DNA-based ConvNet behavior with eight input patterns. **d**, Fluorescence level of each pair of outputs at 24 h after the inputs were added, collected from 88 experiments with 11 example patterns per oracle. Each colored point corresponds to an example pattern, each grey point corresponds to an out-of-class example pattern. The output will be 1 when the fluorescence is greater than the threshold of 0.6, otherwise output will be 0. As expected, all output signals reached ideal 'on' state, while the other fluorescent signals were at ideal 'off' state.

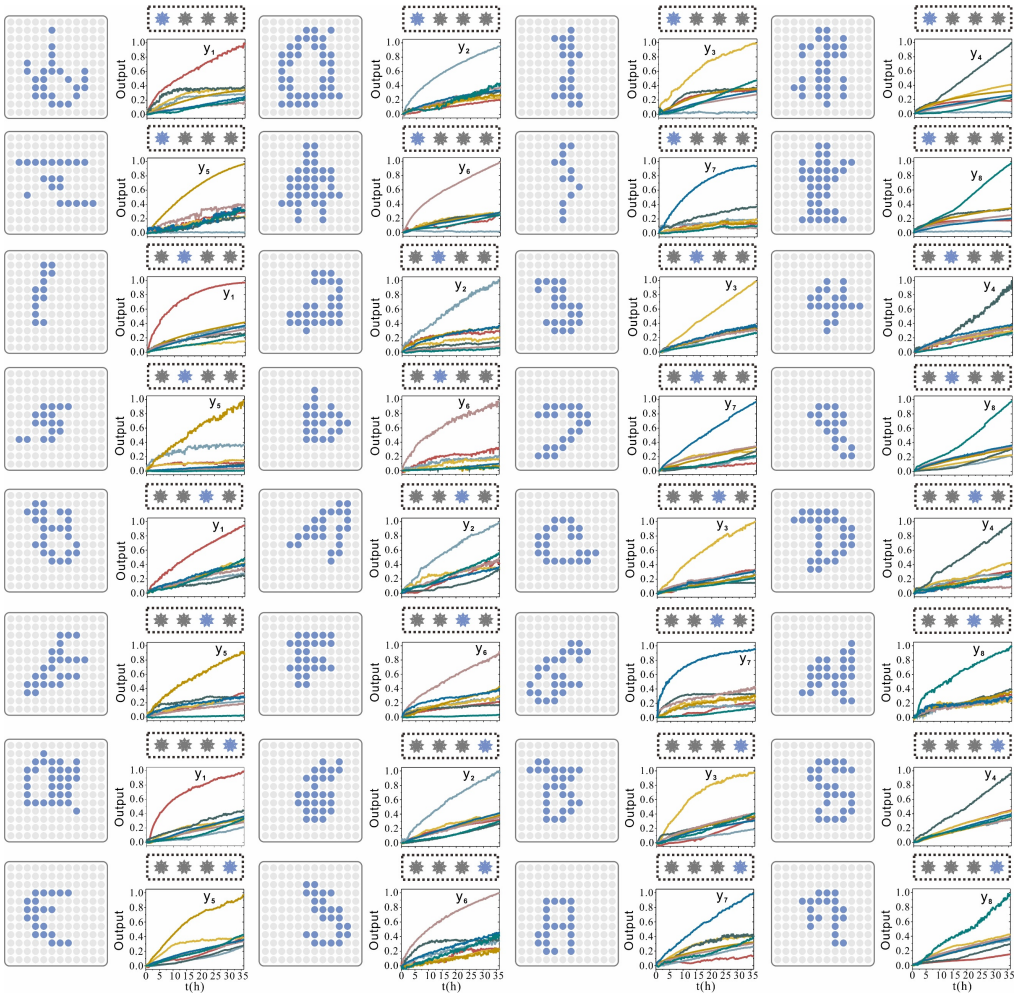
Supplementary Fig. 22| Eleven representative input patterns from the reference dataset.

S7. The DNA-based hierarchical network for one of 32 molecular patterns recognition



Supplementary Fig. 23| 32 patterns were divided into four groups, and each group was trained in silico separately. Once the optimal models were obtained, the value of weights was determined by four convolution kernels.

Supplementary Fig. 24| Fluorescence kinetics data of the circuit behavior of Layer 1 with 32 representative input patterns. Matching colors were used to visualize corresponding trajectories and outputs.



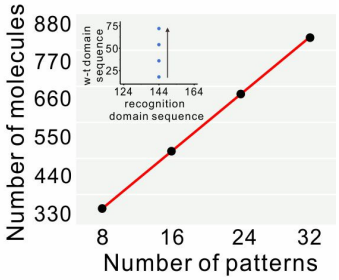
Supplementary Fig. 25| Fluorescence kinetics data of the circuit behavior of Layer 2 with 32 representative input patterns. Matching colors were used to visualize corresponding trajectories and outputs.

a

Reported work	Function	Number of species
Qian et al. (2018)	Pattern recognition	$n\times m+2n+4m+{}^{m}C_{2}$
current work	Pattern recognition	$n+5m+(m+1)\times be$

n, number of bits in a pattern
b, number of groups
m, number of patterns in each group
e, number of bits in a convolution kernel

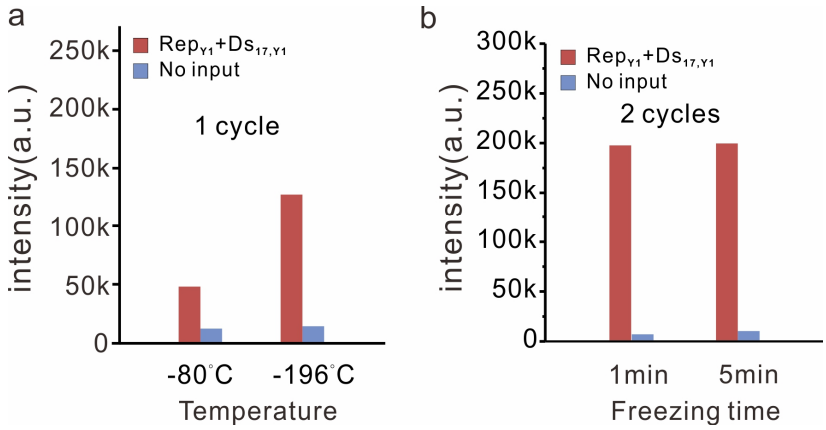
b



Supplementary Fig. 26| Maximum number of species of neural networks implementation for pattern recognition with the two-step classification approach. a, The size of a DNA-based WTA and a DNA-based ConvNet implementation scales up for an increasing number of patterns. In general, constructing a DNA-based ConvNet that can remember b×m distinct n-bit patterns with e-bit kernel size requires n input strands X_i, b×e weight tuning molecules M_{Wt} and b×e×m weight substrate molecules $N_{Wt,Ii,j}$ for weight multiplication, 2m summation gates $Sd_{j,k}$, m complex $Dd_{k,Yi}$ and m annihilation species $Sub_{n,Yi}$ for subtraction, and m reporters Rep_{Yi}, totaling maximum n+5m+(m+1)×b×e molecules. **b**, On the basis of hierarchical network model, the number of molecules used to construct the neural network increased linearly as the number of patterns was scaled up. Moreover, the inset graph shows that the number of w-t (weight-tuning) sequence domains are increased linearly while the number of recognition sequence domains can remain unchanged, which could potentially lead to more scalable and complex DNA sequence design.

S8. Cyclic freeze/thaw approach as drivers of DNA circuits

We discovered that a cyclic physicochemical process (freeze/thaw cycling) can be a potent driver for strand displacement. During freezing, the steep increase in DNA concentrations could trigger the generation of a subpopulation of active complexes; during thawing, the steep drop in DNA concentrations could disassemble active and unproductive complexes. We first sought to study whether a basic strand displacement can be speeded up with freeze/thaw cycling (*Extended data Fig. 5*). Once the invader strand $Ds_{k,Yi}$ was added, the freeze/thaw cycling was completed by two steps (*Extended data Fig. 5a*): (1) the solution was frozen using liquid nitrogen in 1 min. (2) The tube was transferred to a water bath for thawing in 5 mins at 37 °C, then the tube was frozen by liquid nitrogen again. We found that the steady state fluorescence rapidly increased after just one freeze/thaw cycle. For two freeze/thaw cycles of 12 minutes, the extent of reaction was equivalent to that of 15 hours at 25 °C, which corresponded to nearly 75 times increase of the average reaction rate (*Extended data Fig. 5b*). In this case, we can shorten the response time of the reaction from 15 hours to 12 minutes. As the number of cycles increased, the reaction gradually tended to be completed (*Extended data Fig. 5c*). We further explored the effect of freezing temperature and freezing time on DNA strand displacement. The data shows that the speed of DNA strand displacement is faster at -196 °C (*Supplementary Fig. 27a*), and there is no significant increase in speed when increasing freezing time (*Supplementary Fig. 27b*).



Supplementary Fig. 27| a, The fluorescence level of strand displacement with one freeze/thaw cycles at -80°C and -196°C, respectively. **b**, The fluorescence level of strand displacement with two freeze/thaw cycles, with different freezing time.

Reference

1 Phillips, A. Visual DSD: a design and analysis tool for DNA strand displacement systems. *Bioinformatics* **27**, 3211–3213 (2011).