Nonlinear decision-making with enzymatic neural networks

Shu Okumura, Guillaume Gines, Nicolas Lobato-Dauzier, Alexandre Baccouche, Robin Deteix, Teruo Fujii, Yannick Rondelez, Anthony J Genot

Nonlinear decision-making with enzymatic neural networks

S. Okumura[1], G. Gines[2], N. Lobato-Dauzier[1], A. Baccouche[1], R. Deteix[1], T. Fujii[1], Y. Rondelez[2] & A. J. Genot[1*]

1. LIMMS, CNRS-Institute of Industrial Science, IRL 2820, University of Tokyo, 153-8505 Tokyo, Japan. 2. Laboratoire Gulliver, UMR 7083, CNRS, ESPCI Paris, PSL Research University, 10 rue Vauquelin, 75005 Paris, France

Abstract:

Artificial neural networks have revolutionized electronic computing. Similarly, molecular networks with neuromorphic architectures may enable molecular decision-making on a level comparable to gene regulatory networks[1–4]. Nonenzymatic networks could in principle support neuromorphic architectures, and seminal proof-of-principles have been reported[5,6]. However, leakages, as well as issues with sensitivity, speed, nonlinearities and preparation, make the composition of layers delicate, and molecular classifications equivalent to a multilayer neural network (e.g. nonlinear partitioning of a concentration space) remain elusive. Here we introduce DNA-encoded enzymatic neurons with tunable weights and biases, and which are assembled in multilayer architectures to classify nonlinearly separable regions. We first leverage the sharp decision margin of the neurons to compute various majority functions on 10 bits. We then compose neurons into a two-layer network, and synthetize a parametric family of rectangular functions on a microRNA input. Finally, we connect neural and logical computations into a hybrid circuit that recursively partitions a concentration plane according to a decision tree in cell-sized droplets. This computational power and extreme miniaturization open avenues to query and manage molecular systems with complex contents, such as liquid biopsies or DNA databases.

Synthetic DNA has emerged as a versatile polymer to store and process information at the molecular scale. It has powered a rich library of computational molecular devices ranging from logic circuits[7–10] to self-assembling automata[11]. Departing from the biological model of computation, most DNA computing devices imitate the Boolean paradigm of electronics. However, their computing power has fallen short of the exponential growth of Moore's law: their size has been plateauing at ~5-10 logic gates for a decade[8]. In parallel, various groups have started looking at the brain, rather than the CPU, as an inspiration for computing with molecules[2,12,3–6]. This is because neuronal and chemical networks share striking similarities: massively parallel and recurrent architectures, analog and asynchronous operation, fault-tolerant and redundant computations (Figure S12)

In 2018, Lopez et al. reported a DNA-based linear classifier[6] that performs all of its computations with a nonenzymatic mechanism: toehold-mediated strand displacement[13]. Using similar DNA-

only mechanisms on many more inputs and taking inspiration from competitive neural networks[2,3], Cherry and Qian reported in *a tour de force* a DNA classifier for the MNIST database[5]. Together, these molecular classifiers showcased the benefits of neuromorphic networks over Boolean circuits: massive parallelism, handling of analog inputs, and tolerance to corrupted patterns. However, these nonenzymatic classifiers had limited decision margins, *i.e.* they could not discriminate between two similar inputs belonging to different classes. They also suffered from leaks that made the composition of layers delicate. Overall, fully molecular classification was only demonstrated on datasets that could be linearly separated by a wide margin.

DNA-processing enzymes are the workhorse of biotechnology, synthetic biology and molecular biology. They perform an astounding variety of transactions on DNA: production, degradation, cleavage, ligation, scouting, cutting and pasting, or editing. In addition, enzymes are fast and processive, and their kinetic control is tight, making them prime candidates for powering DNA computing devices. Previous reports showcased the power of enzymatic networks for running advanced spatio-temporal dynamics like logic computation[14], switches[15–18], clocks[19,20], predator-prey oscillators[21], quorum sensing[22], spatial waves[23], maze pathfinders[24] or artificial morphogenesis[25,26] - many of which still resist implementation with strand-displacement only. We set out to explore the potential of neuromorphic architectures combining the programmability of DNA with the efficiency of enzymatic processing.

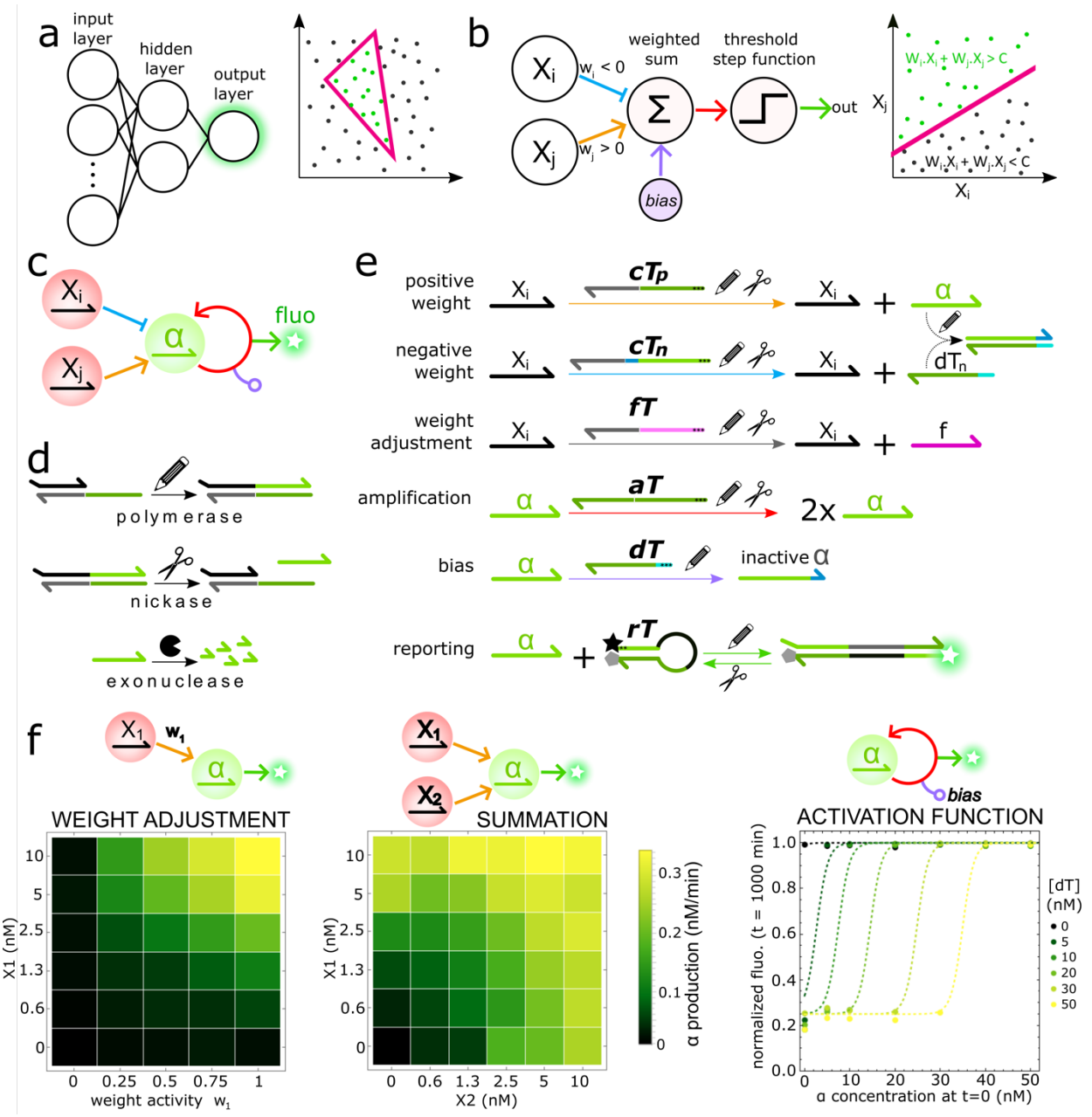


Fig. 1 **Architecture of DNA-encoded enzymatic neural networks. a,** Multilayer neural networks can classify nonlinearly separable regions. **b**, Our individual neuron computes a weighted sum on its inputs and generates an output if the sum exceeds a threshold (linear classification). **c**, Chemical architecture of the neuron. The autocatalytic amplification of the output strand α (red arrow) is triggered when the weighted activation (blue and orange) by input strands X_i and X_j overcomes the thresholding mechanism (purple). **d**, The chemical neuron is powered by three enzymes producing (polymerase), cutting (nickase) and degrading (exonuclease) DNA. **e**, Building blocks of the enzymatic neural networks. Positive and negative weights are computed by converter templates cT_p and cT_n. They produce species α or dT_n whose steady state concentration is proportional to the input X_i. Weights can be independently tuned with fake templates (fT) that compete with cT for the inputs. The activation function -a step function - is encoded in a bistable switch composed of an amplification template (aT, which replicates the species α) and a drain template (dT, which deactivates α and controls the bias). α concentration is monitored using a reporter template (rT). **f**,

Experimental validation of the basic components: weight adjustment on a single input (left), weighted summation on two inputs ($w_1 = 0.5$, $w_2 = 1$) (middle) and application of the step function on α (right). Full traces are available in Extended Data Figure 1.

Linear Classifier

Our neuromorphic networks are built around a generic enzymatic neuron (Fig. 1) that emulates the perceptron proposed by Rosenblatt in 1958[27]. The neuron takes DNA or RNA strands as input. The state of the neuron is encoded by the concentration of a short DNA strand α (the signal sent by the neuron). The neuron computes a weighted sum of its inputs thanks to converter templates. They act like programmable-gain amplifiers in analog electronics, the gain being tuned by the composition of the templates[28] (Fig. 1f). The neuron then takes an ON state (yielding a high concentration of α) if the weighted sum of inputs exceeds a concentration threshold, and remains OFF otherwise (low concentration of α). In modern terms, this mimics a perceptron with a step function as the nonlinear activation function. The inputs here are two DNA analogues of the miR-21 and miR-31 microRNAS (miRNAs), which are involved in cancer[29]. The network is modular and can easily be rewired to accept different inputs or produce new outputs.

We tested the neuron in bulk (~10 µL) (Fig. 2a). It is sensitive (working with subnanomolar inputs), fast (classifying in a few hours) and sharp (OFF and ON states are clearly demarcated). We then shrunk volumes by a factor of 10^5 with droplet microfluidics[30]. We immobilized a layer of droplets with the classifier in a silicon chamber - a material with excellent optical, thermal and mechanical properties[31] - and incubated the chamber in a thermal platform (Extended Data Fig. 2). Overall, this microfluidic setup offers a superb control of concentrations and temperatures, and enables a precise visualization of the decision boundary (Fig. 2 c,d). End-point analysis confirms the exquisite sensitivity and robustness of the classifier in these ~100 pL compartments: OFF and ON regions remain clearly delineated by a linear and sharp boundary (Hill coefficient of ~16, Extended Data Figure 5). This endows the classifier with a narrow decision margin (Fig. 2c): it can discriminate between inputs whose concentrations differ by only ~10-20% (such being the case in majority voting over *10* bits, *vide infra*). The statistical metrics of performance (such as accuracy) are in the range of 80%-99% (Fig.2i)

Versatile linear classification requires negative weights. We opted for a strategy of induced inhibition, in which an input produces a drain template[17], which in turn suppresses the replication of the signal strand - implementing a negative weight (model in Supplementary Information Section S2). This strategy was successful in bulk and droplets (Fig. 2 b,d), also producing a sharp demarcation between OFF and ON, although the boundary is slightly less linear than for positive weights (Hill coefficient of ~63, Extended Data Fig. 5, Supplementary Information S2)

The chemical neuron is analog, and its computation varies continuously with its parameters. On the one hand, this can be used to program the parameters of the neuron, for instance changing the bias by tuning the drain, or the weights by tuning the converters (Fig. 2c). On the other hand, it makes the neuron sensitive to uncertainties on experimental parameters. However, we find that

the deviations are likely to be minimal for a typical operation. More precisely, we analyzed the sensitivity of the bias of a single-input neuron to the drain and temperature. A pipetting error of ~3% on the concentration of drain (typical for a calibrated pipette) translates into a ~10% error on the bias, and an error of ~0.1°C on temperature (typical reproducibility for a thermocycler) translates into a ~2% error on the bias (both measured for drain = 20 nM and T = 41.5°C, Extended Data Figure 4).

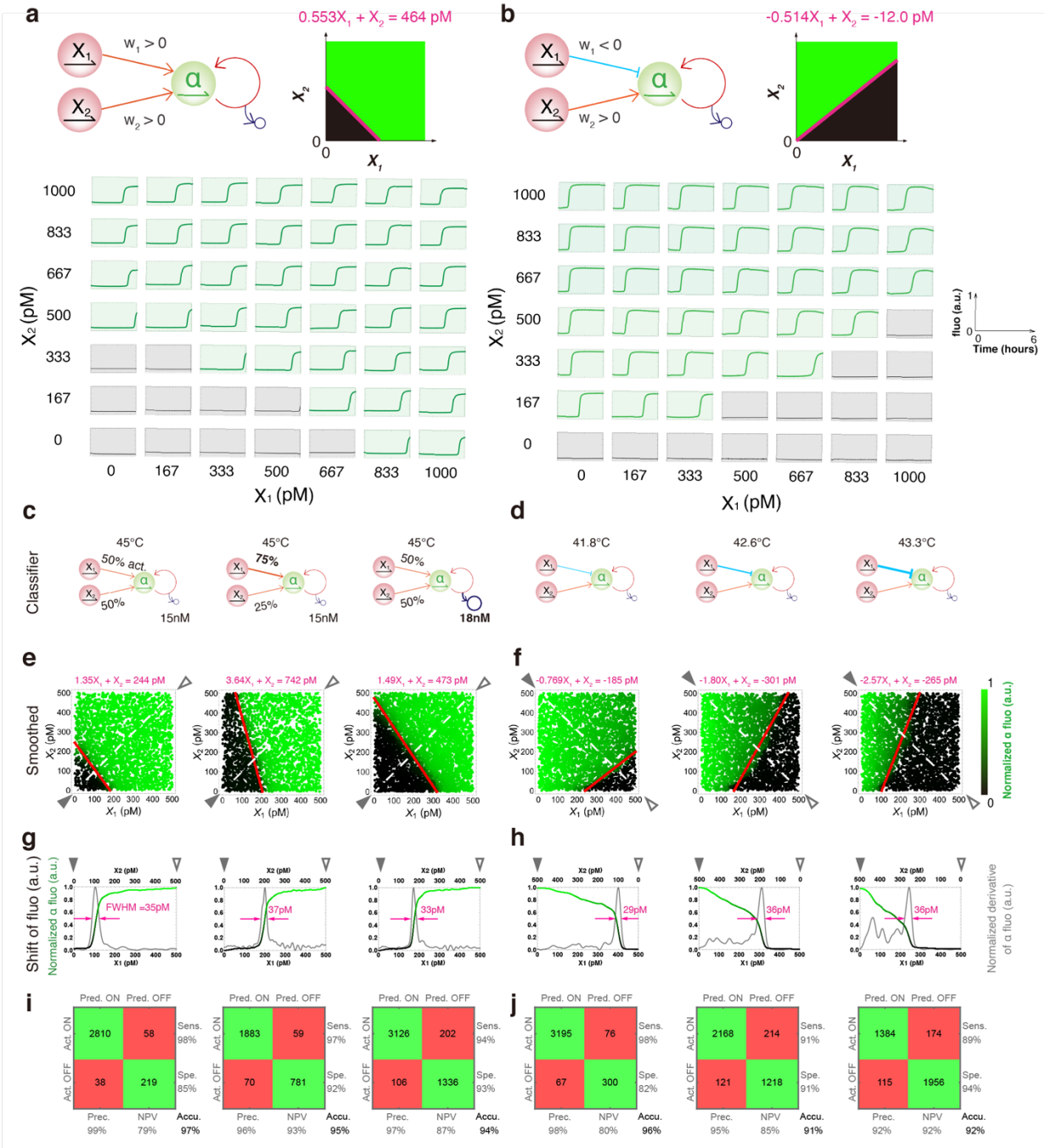


Fig. 2 Operation and tuning of linear classifiers in bulk and microdroplets. Here the input X_1 and X_2 are DNA analogues of miR-21 and miR-31. **a,** Positive-weighted classification with two positive converters

and a bistable switch (left). This classifier partitions the concentration space with a negative-sloped line (right). The matrix of plots shows the dynamics of the classifier in bulk (~10 µl), measured by following the fluorescence of the reporter over 6 hours at 45°C, for varying combinations of inputs. For clarity, the background is green when the classifier finishes in the ON state, and grey otherwise. **b,** Computation of a negative-weighted classifier for the same concentrations of inputs. This classifier partitions the concentration space with a positive-sloped line (right). **c,** Tuning the weights and bias of a positive-weighted classifier (the activity of an input is modulated by changing the composition of its converter and fake templates). **d,** Effect of temperature on a negative-weighted classifier. **e,f** Computation in droplets. The smoothed plots show the fluorescence of the α reporter in droplets prepared with varying concentrations inputs (measured after 6 h). The red line is a linear fit of the OFF/ON boundary. **g,h** Slicing of the fluorescence (black/green curve) and its derivative (grey curve) along one of the diagonal (dashed white line in the smoothed plot). The Full Width at Half Maximum (FWHM) of the derivative is a proxy for the decision margin of classifier (i.e. the distance between unambiguously OFF and ON regions. **i,** The confusion matrices show the number of ON and OFF droplet in each plot, based on their actual (act.) fluorescence, and predicted (pred.) value accordin to the linear fit. The corresponding accuracy, precision, sensitivity, specificity and negative predictive value (NPV) are indicated on the sides of the matrices.

Majority voting

The majority function is a central Boolean primitive essential to many decision rules. It does not have any parameter to tune - a robustness and simplicity which help to overcome a common issue of statistical inference with gene expression datasets: their low sample size and high dimensionality[32]. Majority voting with Boolean DNA circuits has been reported on 3 inputs[33], with designs that do not easily generalize to more bits. Our enzymatic linear classifiers have the necessary decision margin to compute the strict majority function on up to 10 bits (Supplementary Information Section S3). We developed an *ad-hoc* protocol that equalizes the production rate of signal strand for 10 converter templates, in order to ensure the equibalance of each vote (Extended Data Fig. 6). These converters are connected to a neuron whose bias is adjusted just above 5 votes (Fig. 3). We tested the majority network on 10-bits patterns whose number of inputs ranged from 0 to 10, totaling 92 patterns (the two patterns that are fully OFF and fully ON, plus 10 patterns for each number of bits between 1 and 9). The network achieved a classification accuracy of 97% (Fig. 3b,c). Noticeably, the few errors are all localized at the boundary of the decision region (5 inputs). This performance compares favorably with previous nonenzymatic neuromorphic DNA circuits which, based on their reported decision margins, would not be able to perform majority voting beyond 2-3 bits (Supplementary Information Section 3).

We then generalized majority voting with two functionalities: veto and majority right (Supplementary Information Section 3, Figure 3d,e). These rights are arbitrarily granted to any input strand by tuning the nature and concentration of its template. We retested six times the network on 47 patterns, thrice with a veto right given to one of X_3, X_7 or X_8, twice with a majority right to X_3 or X_4, and once without any exception right as a control (Fig. 3b,c). In all cases the accuracy is between 96% and 100%, the only errors clustering again near the decision boundary.

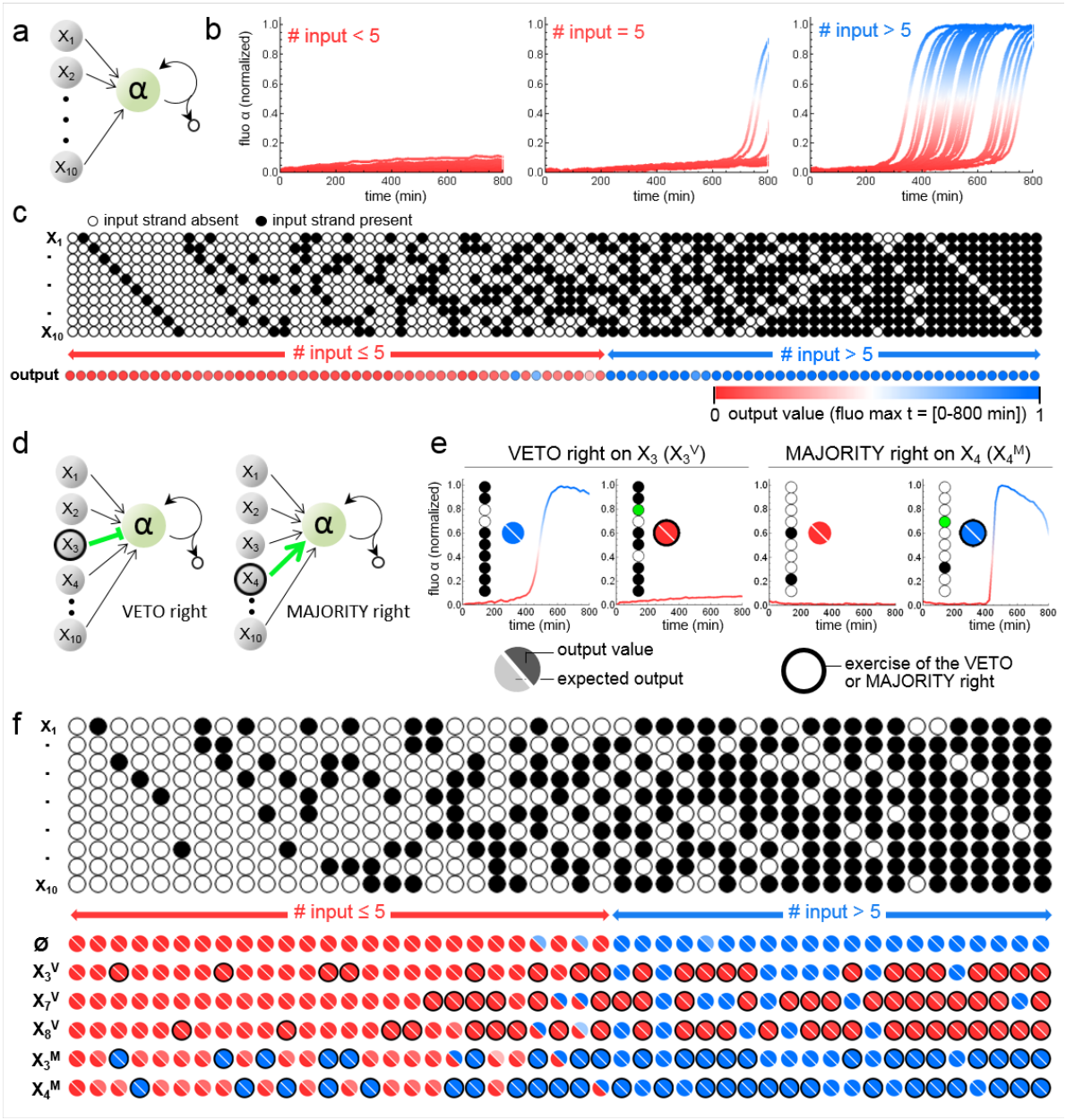


Fig. 3 Majority functions on 10 bits with a high-dimensional linear classifier. **a,** Ten input strands X_i (each input being encoded by a strand at concentration 0 or c) are connected by equal-weight converters to a single neuron. The threshold is adjusted so that more than 5 inputs must be present to trigger the replication of α. (The inputs X_1 and X_2 are unrelated to the strands used in Fig. 1 and Fig. 5) **b,** Bulk time traces of majority voting for 92 inputs patterns. The color of traces reflects the instantaneous level of α. **c,** Summary of results. The output color corresponds to the maximum fluorescence intensity over 800 minutes of incubation. **d,** Majority vote with veto or majority right. A strand is discretionarily granted a veto right by replacing its positive-weighted converter by a negative-weighted converter (veto right). Similarly, a strand is granted a majority right by increasing the concentration of its positive-weighted converter (green arrows). **e**, Examples of majority voting with or without veto right (left) or majority right (right). **f,** Summary of majority voting on a batch of patterns, and with different rules (democratic, veto right to X_3, X_7 or X_8, majority right to X_3 or X_4).

Synthesis of a parametric family of rectangular functions with multilayer networks

Nonlinear classification requires the composition of multiple layers, which we demonstrate here on the canonical example of a rectangular function. Taking a single input at concentration x, this function is constant inside the interval $c_{min}<x<c_{max}$, and null outside. Such window functions are widely used in electronics to filter signals, but biological systems also use them to produce a response when the input is neither too big nor too small[34]. To demonstrate biological relevance, we selected a human microRNA as the input (let7a, involved in development, cancer, aging and metabolism[35]).

To instantiate the two thresholds c_{min} and c_{max} we use a hidden layer with two neurons: α is activated by low concentrations of input, and activates the neuron γ in the second layer, while neuron β is activated at high concentrations of input, but inactivates the neuron γ (Fig. 4a). This architecture – where two neurons have opposite actions on the output - mimics an incoherent feedforward loop (a ubiquitous motif in gene regulatory networks[36]). This three-neurons network defines a family of functions on let7a, which are parametrized by the weights and biases of the neurons. We expect the bias of β – i.e. its drain concentration – to control linearly c_{max}, independently of c_{min}. We thus fixed the concentration of all species but two, the drain for β and the input let7a, scanned them with microfluidics and read out the three neurons at steady state (Fig. 4**b,c**). As expected, the concentration c_{min} is independent of the drain for β (as shown by the vertical black strip in the α plot); the concentration c_{max} varies linearly with the drain for β; and the activated region of the output γ is correctly computed as the intersection of the activated (green) region for α and the inactivated (black) region for β. Taking smoothed horizontal slices of these 2D plots, we extracted the profiles of individual functions in the family (Fig. 4c). Those plots confirm that the output rises and then falls sharply with the input, although the rise is steeper than the fall (this asymmetry is due to the asymmetry of the bistable switch, which is easier to turn ON than OFF). The width of the rectangular function is linear with the drain and varies between ~35 pM and (at least) 90 pM of let7a.

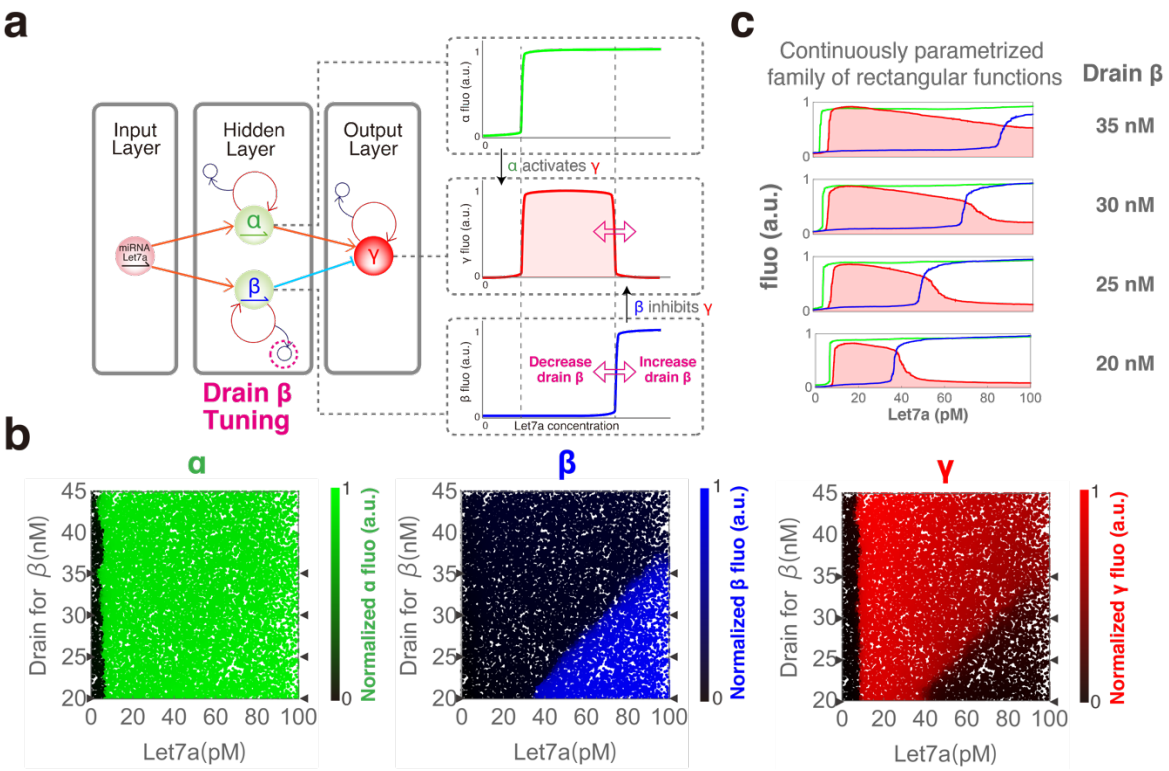


Fig. 4 Synthesis of a parametric family of rectangular functions with a multilayer perceptron. a, A microRNA input activates two neurons (α and β) in the hidden layer. The output neuron γ is activated by α and inhibited by β. As a result of these opposing actions, γ is only active when α is active and β inactive – producing a rectangular function on the input. **b**, Droplet microfluidic scanning of the concentration of input and a parameter of the network (the concentration of drain for β) . The steady state fluorescence of α, β and γ (after 14h) are shown against the concentrations of let7a and β drain. **c**, Smoothed horizontal slices in the γ plot (gray arrows) reveal the tunable rectangular function.

Recursive Space Partitioning by a decision tree

Finally, we composed linear classifiers with a logic gate to classify nonlinearly separable regions and compute a decision tree. In healthcare, decision trees are often used for making a diagnosis based on a clinical presentation, and they are gaining traction in molecular diagnosis, for instance for classifying tumors based on the expression levels of miRNAs[37]. Taking as input a point X = (X_1, X_2) in the concentration plane (Fig. 5a), the algorithm starts from the root node of the tree, and gradually moves toward the leaves. At each node, the algorithm queries the membership of the input to the corresponding half-plane, and moves to either child based on this membership (YES/NO). The algorithm finishes when it reaches a leaf, giving the result of the classification.

Here, we partition the 2D concentration plane into three nonlinearly separable regions (α, β and γ). The network is hybrid and comprises 2 computational layers (Fig. 5 b): a hidden neural layer deciding membership of the α and β region with linear classifiers, and a logical layer deciding membership of the γ region with a NOR gate: its fluorescence is high only when both α and β strands are absent (Extended Data Fig. 7). Membership of the α and β region are computed by two linear classifiers. By tuning the working temperature (Extended Data Fig. 8), we found

conditions where the two linear classifiers become indirectly coupled[21] and α represses β. The network then correctly partitions the concentration space into 3 nonlinearly separable regions (Fig. 5 c,d). We trained two artificial neural networks on the experimental data: a single-layer perceptron and a two-layer perceptron. Unsurprisingly, the single-layer perceptron is unable to correctly classify regions that are nonlinearly separable, while the two-layer perceptron accurately fits our experimental dataset (Fig. 5e)

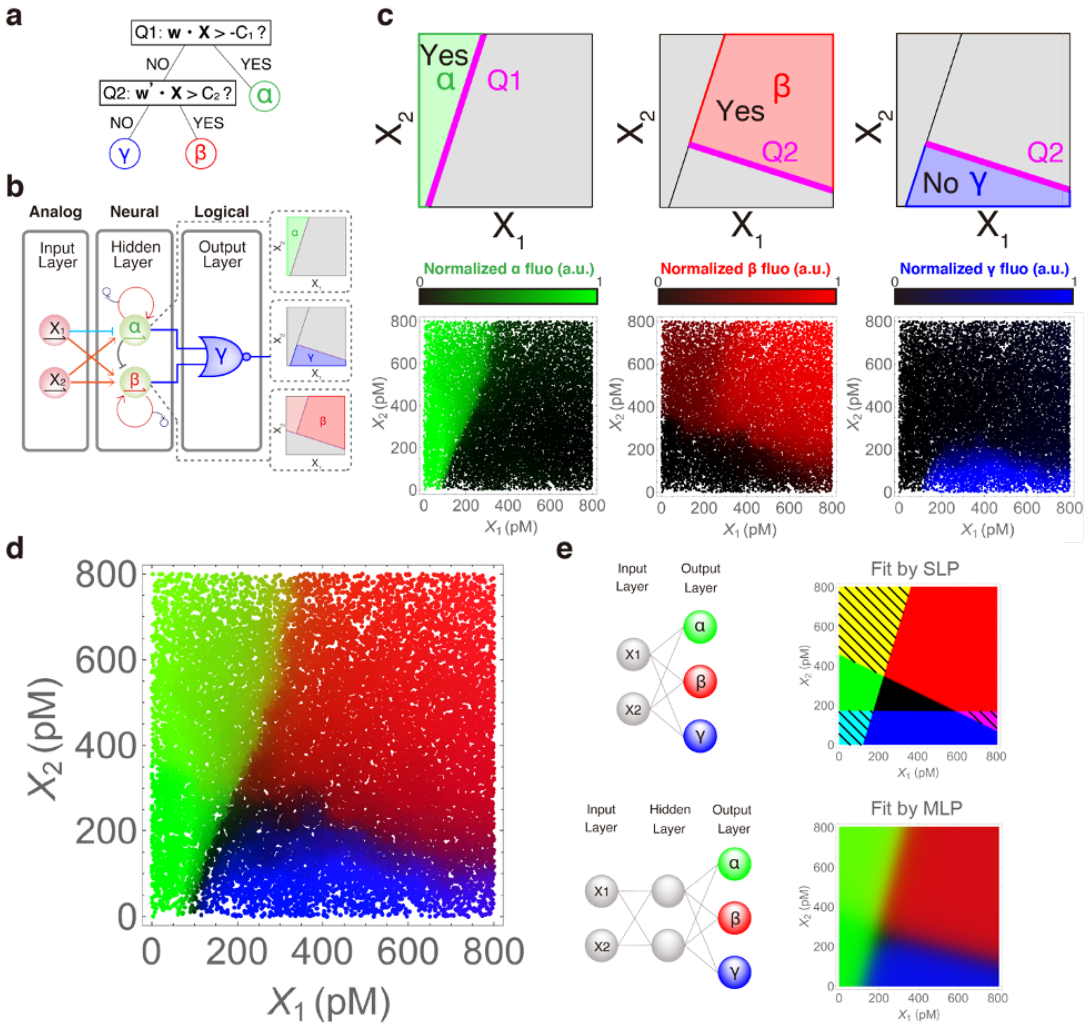


Fig. 5 A hybrid network computing recursive space partitioning. a, A space partitioning tree takes a point X=(X_1, X_2) in the concentration plane, and at each of its nodes, tests if X is a member of the corresponding half-plane. Computation finishes when a leaf is reached. The tree can be traversed in 3 ways, partitioning the plane into three convex regions. **b**, Architecture of a hybrid network computing the partitioning tree (Supplementary Information Fig. S5). The inputs are two strands encoding a position (X_1,X_2) in the concentration plane. The hidden layer is neural and decides membership of the α and β region with two linear classifiers that are indirectly coupled by competitive inhibition (membership is readout by fluorescent reporters). The output layer is logical and decides membership of the γ region with a NOR gate (which turns its fluorescence OFF if X is a member of either α or β region). **c** Fluorescence levels of α, β and γ, measured in ~25,000 droplets after 16 hours. **d,** Merged fluorescence plots **e**, Fit of **d** by a single-layer perceptron (top), and a two-layer perceptron (bottom). The hatched filling indicates erroneous areas where two classes are outputted.

Discussion

Our enzymatic neural networks bring tangible benefits over nonenzymatic ones, namely speed of operation, compactness of network, composition of computations, sharpness of decision margins, sensitivity of detection, correction of errors, and weighing of analog variables with programmable-gain enzymatic amplification (Supplementary Information Section 4). Yet developments will be needed to match the modern form of perceptrons, which apply a wider variety of activation functions than a step function, such as the sigmoid function, for making soft decisions and predicting probabilities[4], or the Rectified Linear Unit (ReLU) to ease training. Sigmoidal responses have long been known in systems biology to be feasible with enzymatic networks (e.g. with cooperativity or push/pull motifs[38]), and simple chemical schemes to compute ReLU were recently proposed[39,40].

More generally, the scale of our networks is sufficient for molecular diagnosis (see below), but work will be needed to reach the scale of *in-silico* machine learning, where neural nets typically have dozens of layers and millions of weights, and are trained on datasets with tens of thousands of examples. In principle, enzymatic networks and droplet microfluidics can handle these scales (as evidenced by the intricate computations performed by gene regulatory networks in cells, or by consortium of single celled organisms), but the current tools for writing, handling, and reading DNA would struggle (Supplementary Information Section 6). However, these hurdles could be overcome by an exponential drop in the cost of DNA synthesis - which is expected in the coming years in response to fields that make heavy use of synthetic DNA. In the long term, enzymatic neural networks could empower the nascent field of DNA data storage[41]. Large amount of data could be stored in DNA databases, and queried, labelled or processed in a massively parallel fashion with enzymatic neural networks and droplet microfluidics.

In the short term, neuromorphic computation could readily find applications in diagnostic. Our enzymatic toolbox previously detected a tumor suppressor miRNA in total RNA from human colon with high specificity and sensitivity[42], and our enzymatic neural networks are similar in size to *in-silico* neural networks that reliably diagnosed breast tumors[43] or prognosed metastasis[44] from multiple molecular clues. This suggests that cancerous patients could be monitored at the point-of-care, using enzymatic neural networks that make diagnosis or prognosis from a panel of miRNAs present in liquid biopsies.

Acknowledgment: This research was supported by the French ANR (grant SmartGuide to A.J.G.), the Japanese JSPS (DC1 fellowship 18J22815 to S.O., postdoctoral fellowship 17F17796 to A.B., and Core-to-Core Program on Advanced Research Networks to T.F., the Japanese MEXT (studentship to N.L.D. and R.D.), the ERC (CoG ProFF 647275 to Y.R. and StG MoP-MiP 949493 to G.G.), and the CNRS (MITI DNA2 grant to A.J.G. and G.G.). We also acknowledge support from the RENATECH microfabrication network in France for the fabrication of the silicon chambers. A.J.G acknowledges support from the ESPCI (Joliot chair), and G.G. acknowledges support from LIMMS for a travel grant. The authors thank Nathanael Aubert-Kato and Leo Cazenille for discussions.

Author contributions: S.O., G.G. and N.L.D. performed the experiments. N.L.D. and R.D. designed the silicon chambers and the thermal gradient platform. A.B. provided support with the enzymatic toolbox. T.F., Y.R., G.G. and A.J.G. supervised the research.

Competing interests T.F.**,** Y.R. and G.G. have filed patents on the PEN DNA toolbox.

Data availability: The data is available upon request

Code availability: The code is available upon request

Methods

1.Materials

DNA strands were purchased from Integrated DNA Technology (IDT) or Biomers (Germany), resuspended in 1X Tris-EDTA buffer (Sigma-Aldrich) and stored at -20 °C. DNA templates were chemically protected from enzymatic degradation with phosphorothioate backbones on their 5' ends.

The DNA polymerase (Vent (exo-), M0257) and nickases (Nb.BsmI,R0706, and Nt.BstNBI, R0607), Bovine Serum Albumin (BSA9000S), were purchased from New England Biolabs, and stored at -20 °C. The exonuclease was expressed by our means in *E. Coli* and purified by chromatography according to a published protocol[45], and stored at -20 °C. This is a thermophile variant of RecJ that works in the temperature range (~40-50 °C) for which the templates were designed, enabling fast melting and release of the output strands from the templates. (Note that that this enzymatic framework can be redesigned to work at 37°C, as demonstrated by recent experiments with biological cells[46])

The surfactant for droplet generation (Fluosurf) was purchased from Emulseo (France) and stored dry at room temperature away from light. Before use, surfactant was freshly dissolved in fluorinated oil (HFE-7500, 3M). Fluorescent dextrans with a molecular weight of 10,000 Da were bought from Thermo Fischer (dextran Cascade Blue D1976, dextran Alexa Fluor 594 D22913 and dextran Alexa Fluor 647 D22914) or Xarxbio (Dextran Cy7 R-CD7-002).

2.Sample preparation

The mastermix (containing all common reagents at constant concentrations and excluding the varying reagents) was assembled on ice. We first mixed the DNA templates, the buffer and the dextrans, then added BSA and enzymes. After gentle vortexing, the master mix was split into several tubes, and the varying reagents were added with their fluorescent dextrans for barcoding.

Bulk fluorescence traces were acquired with a Biorad CFX96 thermocycler.

3.Microdroplet generation

Droplets of varying compositions were generated with an in-house microfluidic platform according to published protocols[30,47]. Briefly, we connected tubes containing an aqueous or oil solution to a flow-focusing microfluidic chip: a PDMS device replicated by soft-lithography from an SU8 or Silicon mold (height of ~55 µm) and plasma bonded to a ~1 mm thick glass slide. The solutions from the aqueous tubes merge inside the chip into a common channel that intersects the oil channel, resulting in the formation of monodisperse droplets at the exit of the junction. The droplets' composition was controlled by varying the pressure applied to each aqueous tube with a pressure controller (MFCS-EZ from Fluigent, France). We scripted the pressure profiles (with AutoIt) to explore an interval, a rectangle, or a cube in the concentration space. The droplets were collected in a pipette tip planted at the outlet before being transferred to a chamber for imaging.

4.Imaging

After droplet generation, the emulsion was spread into a monolayer inside a silicon chamber closed by a coverslip. The silicon chambers were fabricated by standard microfabrication (photolithography and deep Reactive Ion Etching) by opening a square (1 cm x 1 cm) or a rectangle (1 cm x 3 cm) with a depth ~50 µm in a silicon wafer. The coverslip and the chamber were rendered hydrophobic by spin-coating 10% Cytop CTL-809M (Asahi Glass) and baking at 180°C for one hour. After using a pipette to fill the chamber with droplets, the chamber was sealed with a thin coverslip (thickness ~170 µm) by capillarity by leaving a layer of oil between the silicon and the coverslip. Then the chamber was incubated inside a custom platform for thermal control[48] (Extended Data Fig. **2**), which comprised a copper plate (16 cm x 4 cm x 0.5 mm) and two Peltier elements (Adaptive, 40 x 40 mm ET-161-12-08-E) equipped with CPU coolers as a heatsink (Enermax, AM4 ETS – N31 – 02) and controlled by a Peltier controller (TEC-1122, Meerstetter). The temperature near each Peltier element was read by a Pt100 sensor (RS-Pro, 10 mm x 2 mm probe, 4-wire, Class A). The chamber was incubated either in a uniform temperature field (by setting the two Peltier elements to the same temperature), or in a temperature gradient (by setting the two Peltier elements at distinct temperatures). The chamber was imaged with a motorized Nikon Ti2-E epifluorescence microscope, equipped with a LED light source (pE-4000, CoolLed), a sCMOS camera (Prime 95B 25 mm, Photometrics), a 10x objective (CFI Plan Apo Lambda S 10X, NA=0.45, Nikon) and appropriate filters (purchased from Semrock or Chroma). After acquisition, images were unshaded with the BaSiC plugin[49] and stitched[50] in ImageJ before further processing.

5.Analysis

After acquisition, the images were processed with Mathematica according to previous protocols[30,47]. Briefly, the images were segmented to individually detect the droplets, and their fluorescence intensity in each channel was extracted. The fluorescence of the Dextran barcodes was converted into concentrations of inputs or templates, and the fluorescence of the reporters was normalized. For multiplexed experiments, the droplets were separated into subpopulation using the corresponding dextran barcodes. For droplets incubated in a temperature gradient, the local temperature was linearly interpolated with respect to the positions and temperatures of the

Peltier elements. To smooth the raw droplet plots, we replaced the raw fluorescence of each point by the median fluorescence of the *k* nearest points (including the point itself), where *k* is a fraction (typically between 1 and 2% and adjusted according to the plot) of the total number of droplets (which also includes calibration droplets that extends beyond the plotted area[30]). For distributions with dissimilar ranges or physical dimensions (for instance in Figure 3 where the input varies over 100 nM, while the drain varies only over 20 nM, or in Extended Data Figure 4 where one coordinate is a concentration while the other is a temperature), we normalized the distributions by their standard deviation before finding the nearest neighbors.

We fitted the separatrix of the linear classifiers as follows. First, we bin the droplets along one axis, say X_1, by groups of ~100 droplets. For each bin, we then determine the position of the boundary along the X_2 axis by running a moving median on the fluorescence of the droplets, and finding the first X_2 for which the median hits a given threshold. (We chose a common threshold of 0.3, rather 0.5, because it fitted better the datasets for the negative classifier – which had a slightly lower fluorescence at their boundaries). This procedure yielded a set of points on the boundary, which we linearly fit to extract the equation of the separatrix.

We used Mathematica to train a single-layer perceptron (3 nodes in its output layer) and a multi-layer perceptron (2 nodes in its hidden layer, 3 nodes in its output layer) on the smoothed data of the space partitioning network. The activation function is a logistic sigmoid. For the training of the SLP, we adjusted the initial weights and biases to approximately match the boundaries of the 3 regions, which was found to improve the subsequent training. We trained both nets for 50 rounds, without a test set.

1. Hjelmfelt, A., Weinberger, E. D. & Ross, J. Chemical implementation of neural networks and Turing machines. *Proc. Natl. Acad. Sci.* **88**, 10983–10987 (1991).
2. Kim, J., Hopfield, J. J. & Winfree, E. Neural network computation by in vitro transcriptional circuits. *Adv. Neural Inf. Process. Syst.* **17**, 681–688 (2004).
3. Genot, A. J., Fujii, T. & Rondelez, Y. Scaling down DNA circuits with competitive neural networks. *J. R. Soc. Interface* **10**, 20130212 (2013).
4. Salehi, S. A., Liu, X., Riedel, M. D. & Parhi, K. K. Computing Mathematical Functions using DNA via Fractional Coding. *Sci. Rep.* **8**, 8312 (2018).
5. Cherry, K. M. & Qian, L. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **559**, 370–376 (2018).
6. Lopez, R., Wang, R. & Seelig, G. A molecular multi-gene classifier for disease diagnostics. *Nat. Chem.* **10**, 746–754 (2018).

7. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-Free Nucleic Acid Logic Circuits. *Science* **314**, 1585–1588 (2006).
8. Qian, L. & Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **332**, 1196–1201 (2011).
9. Genot, A. J., Bath, J. & Turberfield, A. J. Reversible Logic Circuits Made of DNA. *J. Am. Chem. Soc.* **133**, 20080–20083 (2011).
10. Wang, F. *et al.* Implementing digital computing with DNA-based switching circuits. *Nat. Commun.* **11**, 121 (2020).
11. Woods, D. *et al.* Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. *Nature* **567**, 366–372 (2019).
12. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011).
13. Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605–608 (2000).
14. Song, T. *et al.* Fast and compact DNA logic circuits based on single-stranded gates using strand-displacing polymerase. *Nat. Nanotechnol.* **14**, 1075–1081 (2019).
15. Kim, J., White, K. S. & Winfree, E. Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Mol. Syst. Biol.* **2**, 68 (2006).
16. Padirac, A., Fujii, T. & Rondelez, Y. Bottom-up construction of in vitro switchable memories. *Proc. Natl. Acad. Sci.* **109**, E3212–E3220 (2012).
17. Montagne, K., Gines, G., Fujii, T. & Rondelez, Y. Boosting functionality of synthetic DNA circuits with tailored deactivation. *Nat. Commun.* **7**, 13474 (2016).
18. Meijer, L. H. H. *et al.* Hierarchical control of enzymatic actuators using DNA-based switchable memories. *Nat. Commun.* **8**, 1117 (2017).
19. Franco, E. *et al.* Timing molecular motion and production with a synthetic transcriptional clock. *Proc. Natl. Acad. Sci.* **108**, E784–E793 (2011).

20. Kim, J. & Winfree, E. Synthetic in vitro transcriptional oscillators. *Mol. Syst. Biol.* **7**, 465 (2011).

21. Fujii, T. & Rondelez, Y. Predator–Prey Molecular Ecosystems. *ACS Nano* **7**, 27–34 (2013).

22. Gines, G. *et al.* Microscopic agents programmed by DNA circuits. *Nat. Nanotechnol.* **12**, 351–359 (2017).

23. Padirac, A., Fujii, T., Estévez-Torres, A. & Rondelez, Y. Spatial Waves in Synthetic Biochemical Networks. *J. Am. Chem. Soc.* **135**, 14586–14592 (2013).

24. Zambrano, A., Zadorin, A. S., Rondelez, Y., Estévez-Torres, A. & Galas, J.-C. Pursuit-and-Evasion Reaction-Diffusion Waves in Microreactors with Tailored Geometry. *J. Phys. Chem. B* **119**, 5349–5355 (2015).

25. Zadorin, A. S. *et al.* Synthesis and materialization of a reaction–diffusion French flag pattern. *Nat. Chem.* **9**, 990–996 (2017).

26. Dupin, A. & Simmel, F. C. Signalling and differentiation in emulsion-based multi-compartmentalized in vitro gene circuits. *Nat. Chem.* **11**, 32–39 (2019).

27. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**, 386–408 (1958).

28. Wilhelm, D., Bruck, J. & Qian, L. Probabilistic switching circuits in DNA. *Proc. Natl. Acad. Sci.* **115**, 903–908 (2018).

29. Slaby, O. *et al.* Altered Expression of miR-21, miR-31, miR-143 and miR-145 Is Related to Clinicopathologic Features of Colorectal Cancer. *Oncology* **72**, 397–402 (2007).

30. Genot, A. J. *et al.* High-resolution mapping of bifurcations in nonlinear biochemical circuits. *Nat. Chem.* **8**, 760–767 (2016).

31. Lobato-Dauzier, N. *et al.* Silicon chambers for enhanced-imaging of droplet arrays in a graded temperature field. in *microTAS 2019* (2019).

32. Yata, K. & Aoshima, M. Effective PCA for high-dimension, low-sample-size data with noise reduction via geometric representations. *J. Multivar. Anal.* **105**, 193–215 (2012).

33. Li, W., Yang, Y., Yan, H. & Liu, Y. Three-Input Majority Logic Gate and Multiple Input Logic Circuit Based on DNA Strand Displacement. *Nano Lett.* **13**, 2980–2988 (2013).

34. Cataudella, I., Sneppen, K., Gerdes, K. & Mitarai, N. Conditional Cooperativity of Toxin - Antitoxin Regulation Can Mediate Bistability between Growth and Dormancy. *PLOS Comput. Biol.* **9**, e1003174 (2013).

35. Su, J.-L., Chen, P.-S., Johansson, G. & Kuo, M.-L. Function and regulation of let-7 family microRNAs. *MicroRNA Shariqah United Arab Emir.* **1**, 34–39 (2012).

36. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci.* **100**, 11980–11985 (2003).

37. Zakrzewska, M. *et al.* Expression-based decision tree model reveals distinct microRNA expression pattern in pediatric neuronal and mixed neuronal-glial tumors. *BMC Cancer* **19**, 544 (2019).

38. Goldbeter, A. & Koshland, D. E. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci.* **78**, 6840–6844 (1981).

39. Moorman, A., Samaniego, C. C., Maley, C. & Weiss, R. A Dynamical Biomolecular Neural Network. in *2019 IEEE 58th Conference on Decision and Control (CDC)* 1797–1802 (2019). doi:10.1109/CDC40024.2019.9030122.

40. Vasic, M., Chalk, C., Khurshid, S. & Soloveichik, D. Deep Molecular Programming: A Natural Implementation of Binary-Weight ReLU Neural Networks. in *International Conference on Machine Learning* 9701–9711 (PMLR, 2020).

41. Hao, Y., Li, Q., Fan, C. & Wang, F. Data Storage Based on DNA. *Small Struct.* **2**, 2000046 (2021).

42. Gines, G. *et al.* Isothermal digital detection of microRNAs using background-free molecular circuit. *Sci. Adv.* **6**, eaay5952 (2020).

43. McDermott, A. M. *et al.* Identification and Validation of Oncologic miRNA Biomarkers for Luminal A-like Breast Cancer. *PLOS ONE* **9**, e87032 (2014).

44. Lancashire, L. J. *et al.* A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Res. Treat.* **120**, 83–93 (2010).

45. Yamagata, A., Masui, R., Kakuta, Y., Kuramitsu, S. & Fukuyama, K. Overexpression, purification and characterization of RecJ protein from Thermus thermophilus HB8 and its core domain. *Nucleic Acids Res.* **29**, 4617–4624 (2001).

46. Van Der Hofstadt, M., Galas, J.-C. & Estevez-Torres, A. Spatiotemporal Patterning of Living Cells with Extracellular DNA Programs. *ACS Nano* **15**, 1741–1752 (2021).

47. Baccouche, A. *et al.* Massively parallel and multiparameter titration of biochemical assays with droplet microfluidics. *Nat. Protoc.* **12**, 1912–1932 (2017).

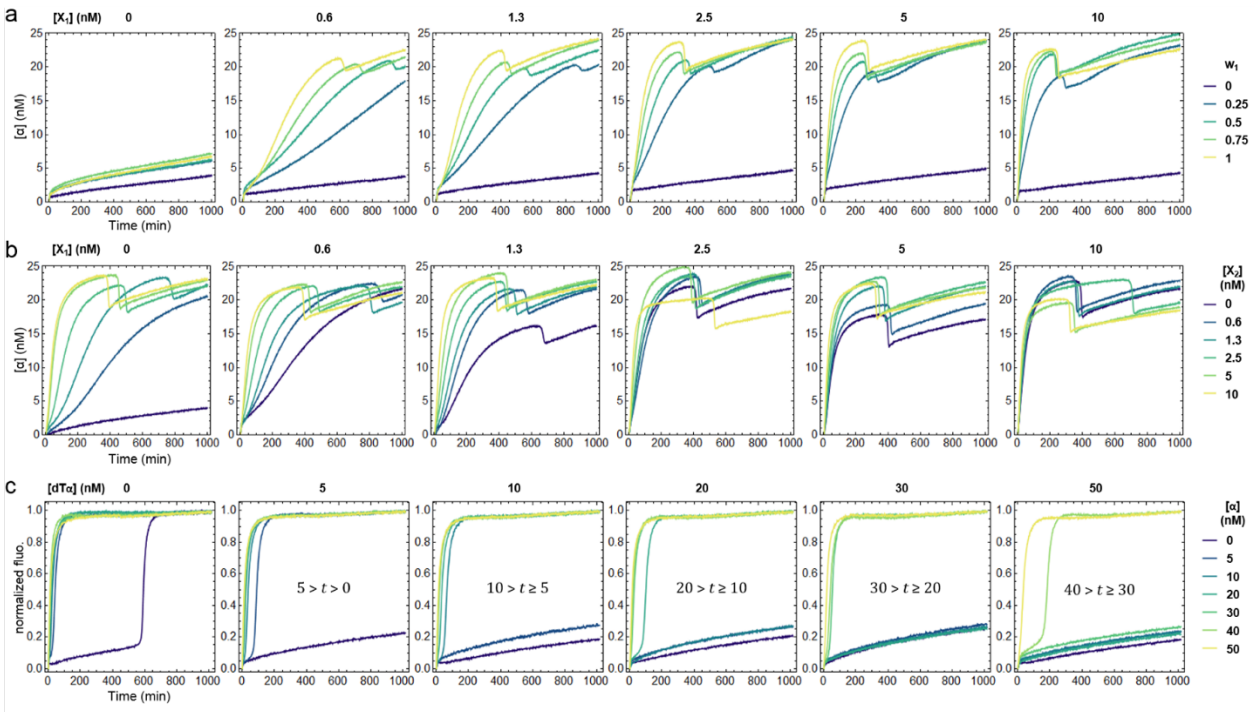
48. Deteix, R. *et al.* Droplet-based investigation of a biochemical bistable circuit for sensitive and noise-free detection of nucleic acid. in *microTAS2019*.

49. Peng, T. *et al.* A BaSiC tool for background and shading correction of optical microscopy images. *Nat. Commun.* **8**, 14836 (2017).

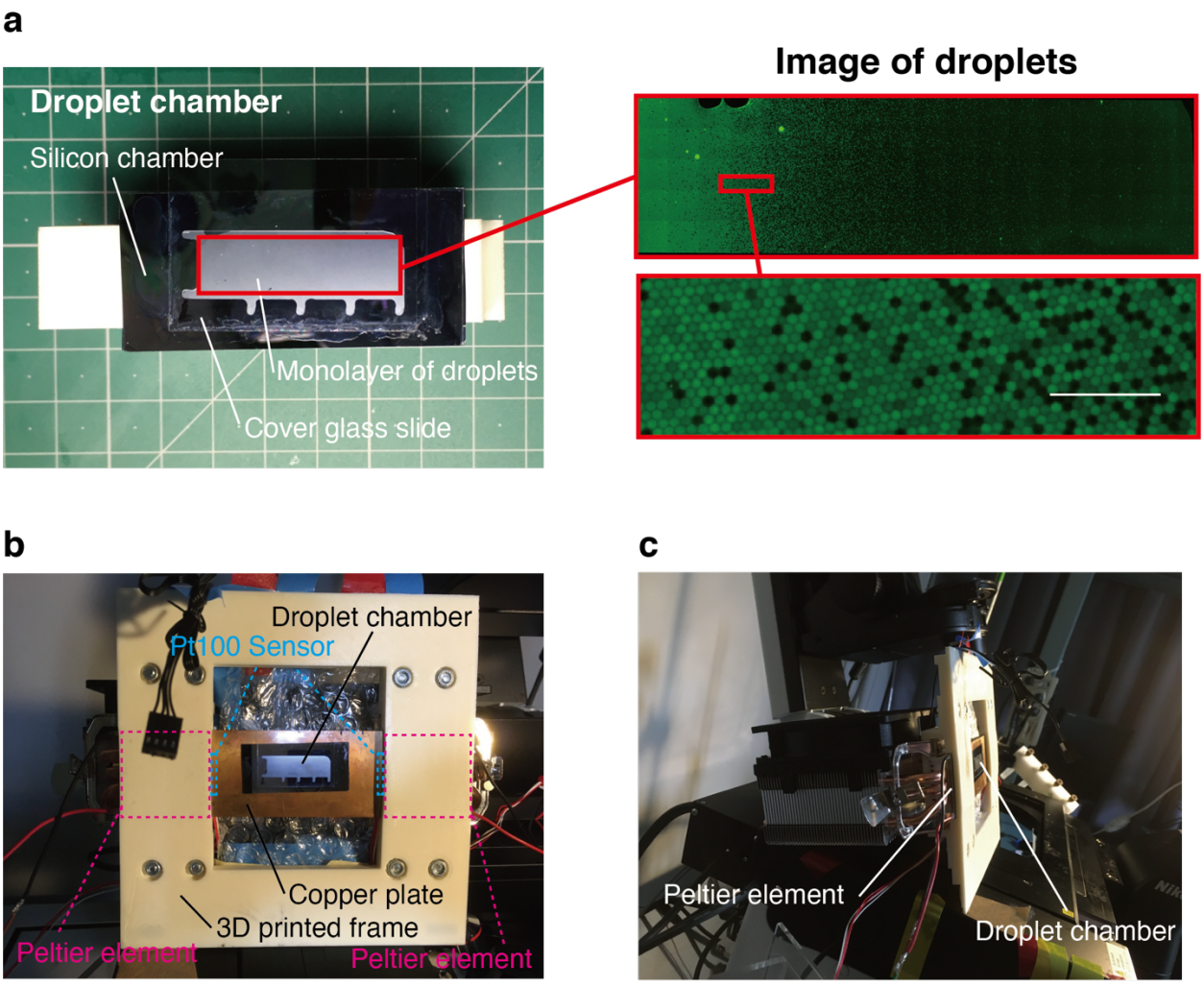
50. Preibisch, S., Saalfeld, S. & Tomancak, P. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* **25**, 1463–1465 (2009).

****

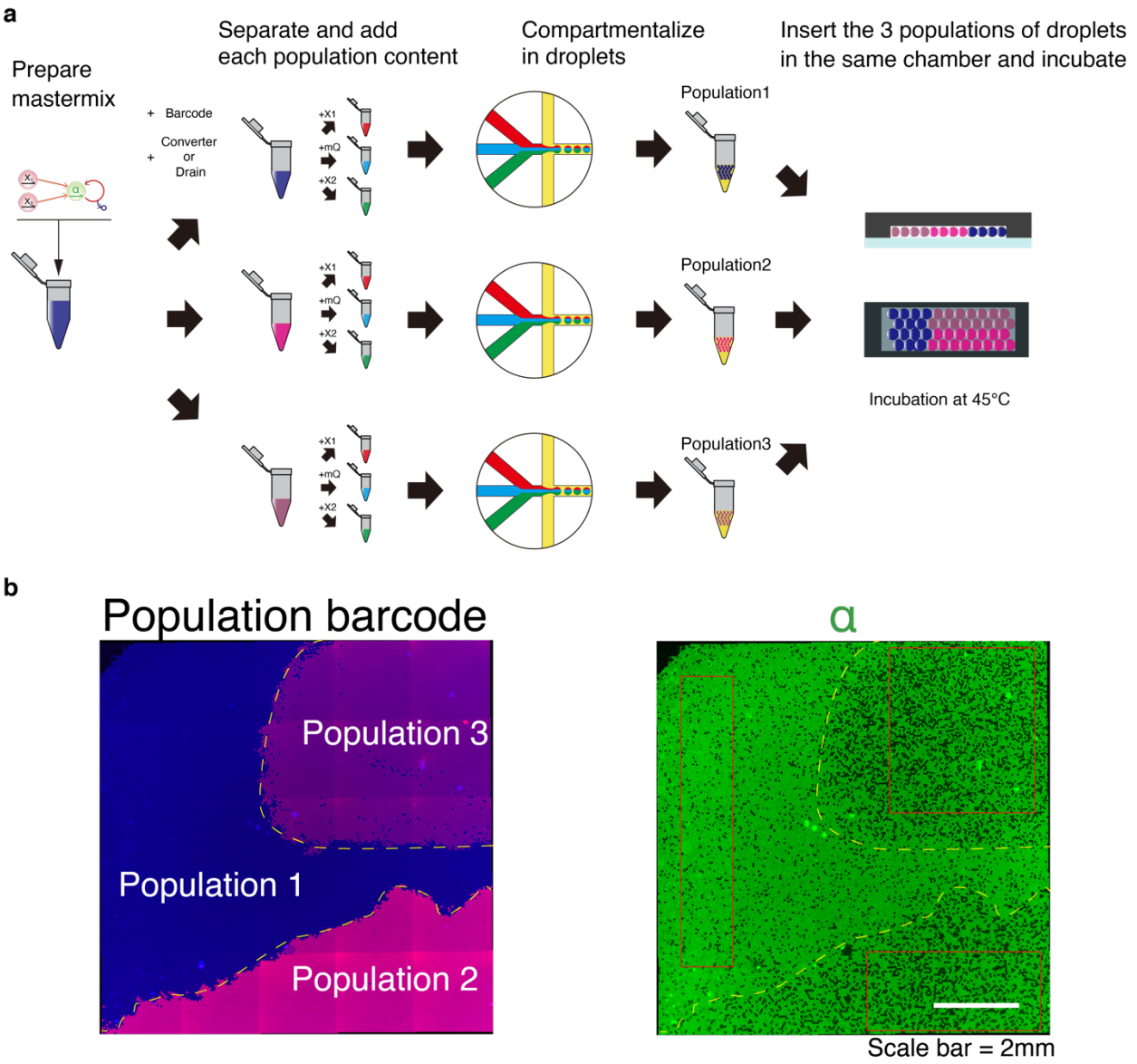
Extended Data



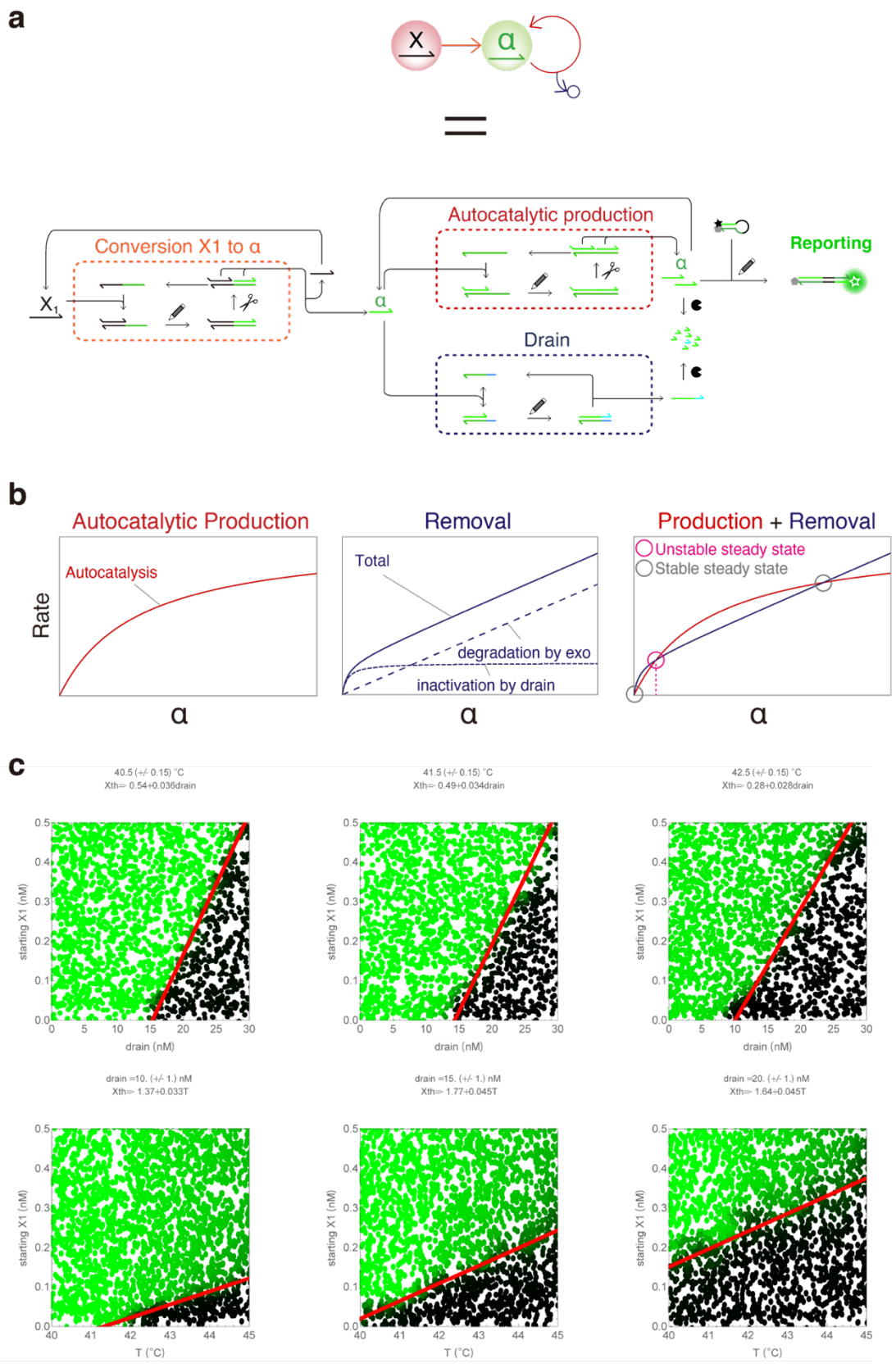
Extended Data Fig.1 Experimental validation of the basic components of a chemical neuron. To demonstrate the weighted summation mechanism (**a**, weight adjustment and **b**, summation), we used high concentrations of input strands, which allows for the direct visualization of the linear production of α strand (i.e. in absence of the amplification reaction that composes the activation function). The threshold activation function and the possibility to control the bias (**c**) is estimated by measuring the concentration of α required to trigger the amplification reaction at given drain template (dTα) concentrations. **a**, Weight adjustment: w_1 corresponds to the ratio of X_1toα (cT) and X_1tof (fT) (10 nM total). The production of α from various concentrations of X_1 is directly monitored using 25 nM of rTα. **b**, Summation of X_1 and X_2: all samples contain 5 nM X_1toα, 5 nM X_1tof (w_1 = 0.5) and 10 nM X_2toα (w_2 = 1). The production of α from different concentrations of X_1 and X_2 is directly monitored using 25 nM of rTα. **c**, Activation function: the amplification reaction of samples containing various initial concentrations of α is monitored in real-time. The bias (i.e. amplification threshold, noted *t*) is tuned according to the concentration of dTα.



Extended Data Fig. 2 Thermal setup for droplet incubation and imaging. a, After generation, an emulsion is spread into a monolayer of droplets inside a silicon chamber. Silicon offers ideal conditions for imaging and incubation: high thermal conductivity, mechanical rigidity and optical reflectivity. The inset shows a monolayer of droplets imaged by fluorescence. Scale bar = 500 µm. **b,** Bottom view of the setup. The chamber is fixed against a copper plate by capillarity by sandwiching a drop of mineral oil between the plate and the chamber. Two Peltier elements, separated by ~ 7.5 cm, impose a thermal gradient across the copper plate. Pt100 sensors report the local temperature near each Peltier element to a Peltier controller. The whole setup is encased into a 3D printed frame that is fitted into the microscope stage. **c,** Side view of the setup. Heat is extracted from the Peltier elements with CPU cooling fans.

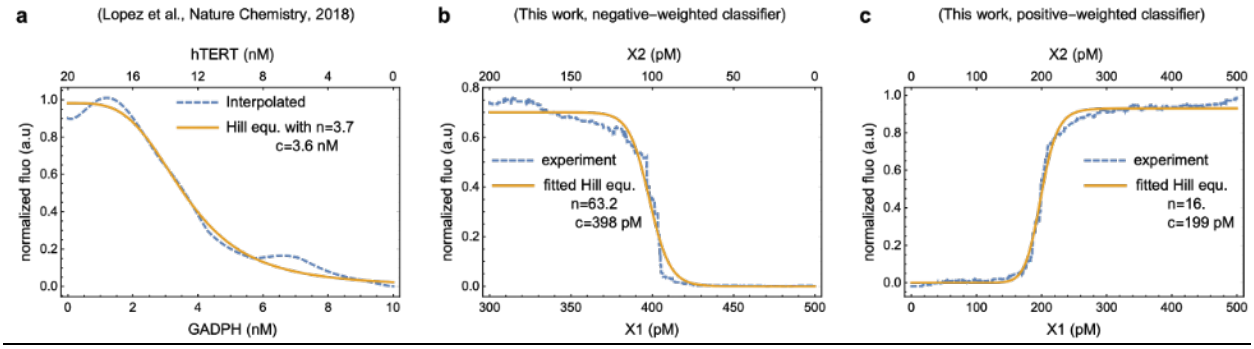


Extended Data Fig. 3 Microfluidic workflow for measuring the dependence of the separatrix on the concentration of reagents. a, We determined the equation of the separatrix in the (X_1,X_2) plane for 3 conditions in a multiplexed experiment. We first performed 3 rounds of droplet generation. During each round, we scanned the (X_1,X_2) plane by mixing 3 tubes (which all derive from the same master mix) and collected the emulsion in a separate tube. Between each round, we changed the set of 3 tubes - thus changing their common mastermix. This allowed us to vary the concentrations of converter templates or drain templates in the master mix. After generation, we simultaneously imaged the three subpopulations together. To that end, we sequentially and gently filled a chamber with each population: the subpopulations remained spatially separated. We then incubated and imaged the chamber. During image processing, we separated the subpopulation by selecting 3 distinct regions (shown as red boxes in the right picture).

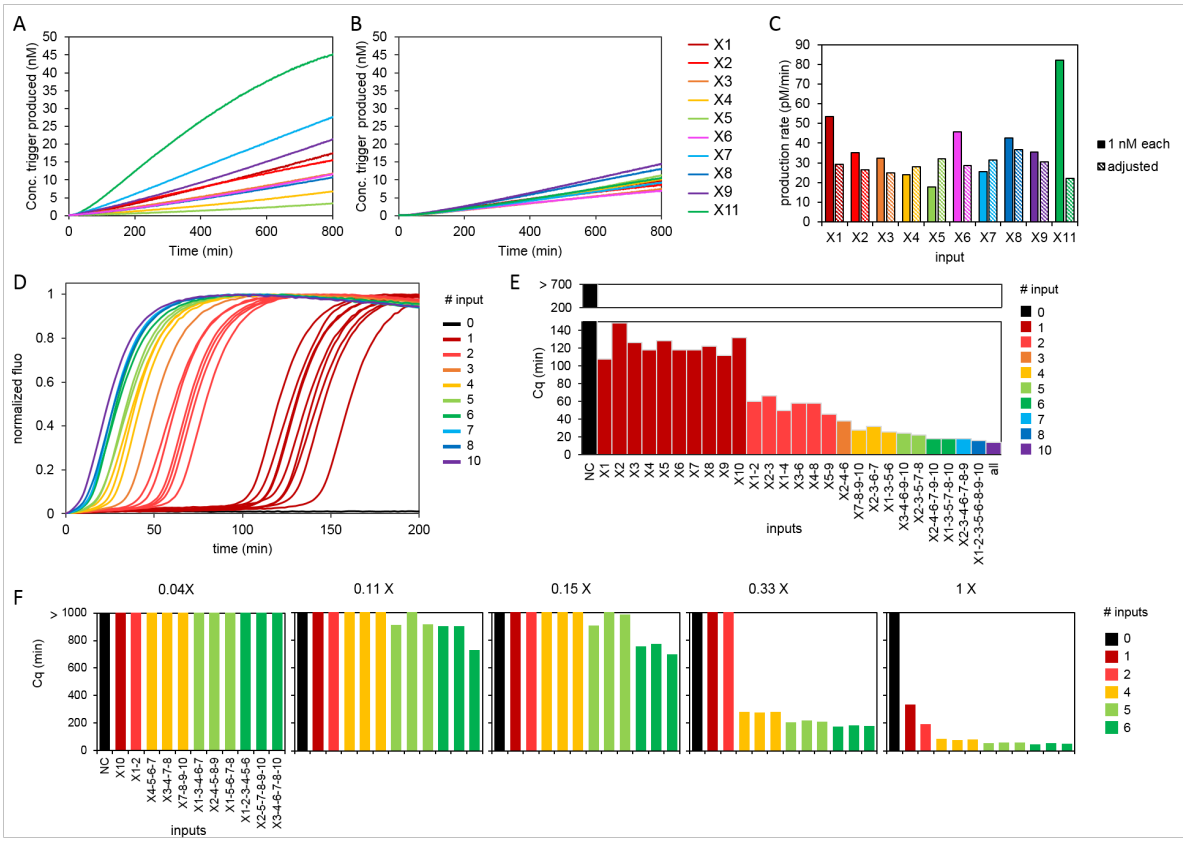


Extended Data Fig.4 See next page for caption.

Extended Data Fig. 4 Design and kinetics of the linear classifier. a, Full architecture of the linear classifier for one input X_1. The classifier comprises three templates: a converter template (which produces α when bound to its input X_1), a autocatalytic template (which autocatalytically replicates α), and a drain template (which deactivates α). The strand α is continuously degraded by the exonuclease. **b,** Production and removal curves showing the rate of production of α by the autocatalytic template (left), and the rate of removal of α by the drain template and the exonuclease (middle). Inactivation by drain template is fast but quickly saturated, while degradation by the exonuclease is slow but linear. This interplay creates a kink in the removal curve, resulting in the existence of 3 intersection points between the production curve and the removal curve. The upper and lower intersection points are the two stable steady states (OFF and ON), while the middle point is an unstable steady state, a threshold which controls the crossover between autocatalytic production and removal. If α is over the threshold, the drain is saturated, production overcomes the removal by the exonuclease and the drain, and α is amplified up to the ON state. Otherwise, α is removed down to the OFF state. The existence of an unstable steady state is controlled by the shape of production and removal curves, which must intersect at three points. The shape of the removal curve is controlled by the concentrations of drain templates, exonuclease and polymerase (for the inactivation step in the drain). The shape of the production curve is controlled by the concentration of autocatalytic templates, polymerase and nickase. **c,** Microfluidic mapping of the dependence of the bias on X1 to the drain. We prepared droplets with varying concentrations of drain for α and input X_1. We incubated the droplets in a temperature gradient, and imaged their content after 6 hours. The top plots show the fluorescence of droplets in the space (drain, X_1), the colour indicating the level of α. The bottom plots show the fluorescence of droplets in the space (temperature, X_1). The red lines are linear fit of the boundary, with equation indicated above each plot. The concentration are in nM and the temperature in Celsius.

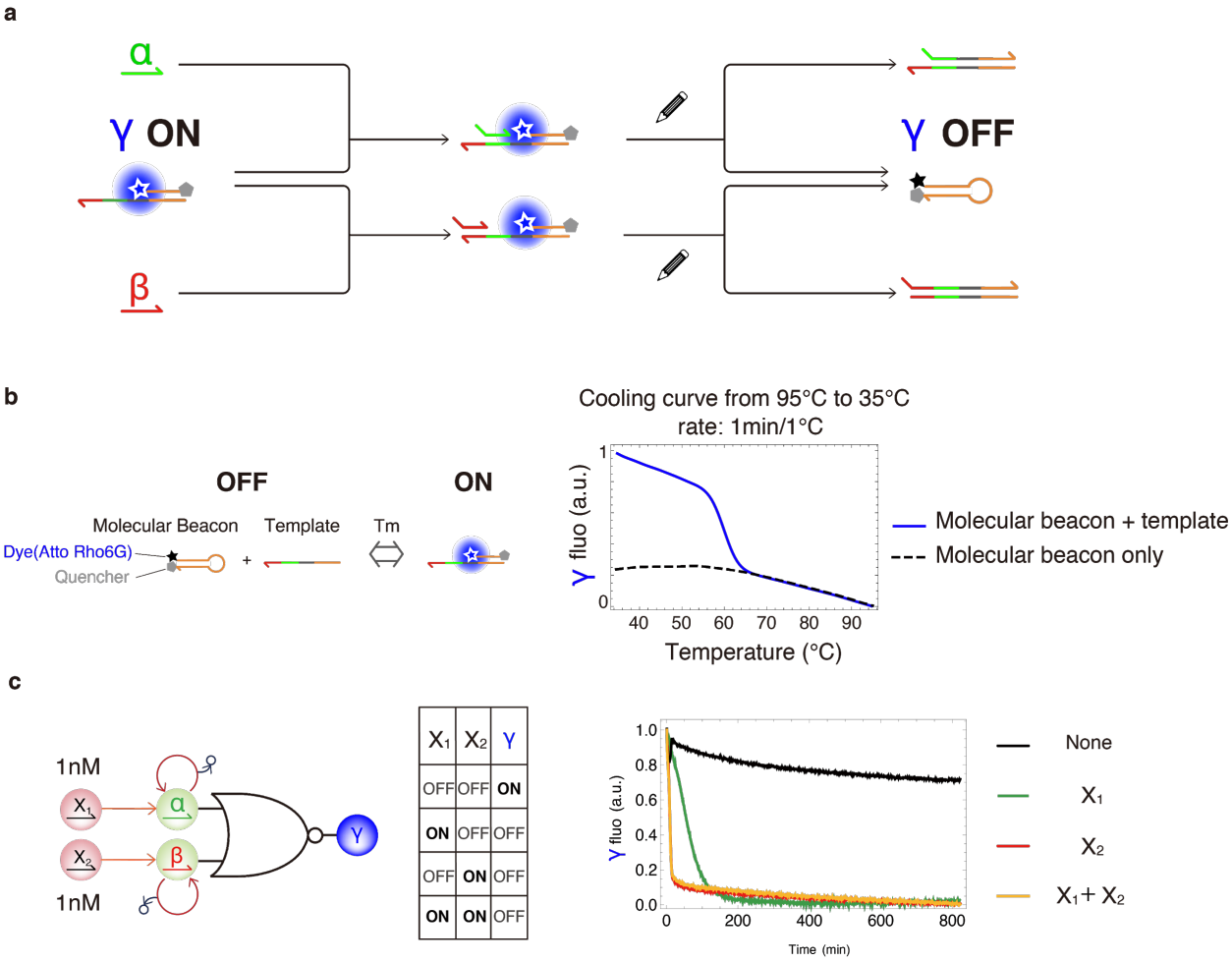


Extended Data Fig. 5 Enzymatic classifiers have stronger nonlinearities and higher sensitivities than a state-of-the-art nonenzymatic classifier. a, Steady-state fluorescence of a reported negative-weighted nonenzymatic linear classifier[6]. The fitted Hill equation $c^n/(c^n+x^n)$ is shown in plain, the interpolation from the 8 points along the diagonal of Fig. 3.d of reference[6] is shown in dashed. **b,** Steady-state fluorescence of our negative-weighted enzymatic linear classifier (Fig. 2d, left), and its fit by a Hill equation $c^n/(c^n+x^n)$. **c,** Steady-state fluorescence of our positive-weighted enzymatic linear classifier (Fig.2c, middle) and its fit by a Hill equation $x^n/(c^n+x^n)$. Hill equations were fitted with a prefactor to improve the goodness of fit.

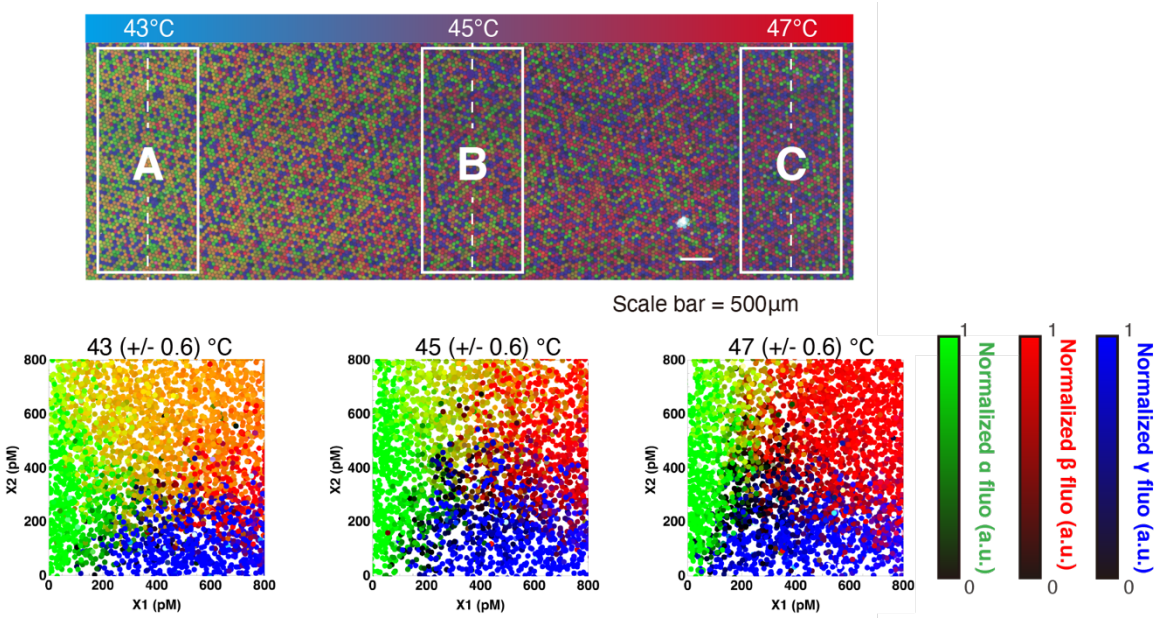


Extended Data Fig. 6 Majority voting weight adjustment. a-b. Trigger production by 10 converter templates at 1 nM concentration (a) or adjusted proportionally to the production rate, arbitrarily choosing X3 as reference (b). **c**. Comparison of the production rate from panel a. and b. At constant cT concentration, we observed large rate discrepancies depending on the input sequence (mean production rate = 39 pM ± 18 pM/min). A factor of 4.6 was computed between the fastest and the slowest converter template. Balancing the cT concentration reduces the coefficient of variation from 47 % to 14 % (29 +/- 4 pM/min). d-e. Majority voting with balanced cT concentrations. In a first attempt, we perform a majority voting experiment on 28 samples. Concentrations of the 1X cT bundle were 0.8, 0.7, 1.2, 1.9, 3, 0.6, 1.2, 0.5, 1.4, 1.4 nM, respectively from X1 to X10. Enzyme concentrations were set to 70 u/mL Vent(exo-), 300 u/mL Nb.BsmI and 23 nM of ttRecJ. **d**. Amplification curves for the 28 samples spiked with 0 to 10 inputs (5 nM each, various combinations). **e**. Bar chart of the amplification times (Cq). As expected, Cq are negatively correlated to the number of spiked inputs (the more inputs, the faster the amplification). Interestingly, all ten inputs exert a consistent activation force on the switch, triggering the amplification between 112 and 148 minutes when spiked individually (123±11 min on average, hence a coefficient of variation of less than 10 %). In these conditions, all samples amplify within 150 minutes (except for the negative control, NC), suggesting that the production rate for each input is too high to unambiguously classify samples with less than 5 inputs from samples with more than 5 inputs. **f**. Effect of the cT bundle concentration on sample classification. To further reduce the weight of all inputs, we decreased the concentration of all converter templates from 1 X to 0.04 X. For the lowest concentration of cT bundle, no amplification is observed within 1000 minutes for all samples (from 0 to 6 inputs). 1X cT bundle results reproducibly in the amplification of all samples (except the negative control, NC), with a poor discrimination between 4, 5 and 6 inputs samples. Interestingly, we observed a sharp threshold between ≤ 2 and > 2 inputs with a 3-fold dilution of the cT bundle concentration. This demonstrates that the amplification threshold can be tuned finely by adjusting the production rate of all inputs, allowing to set an arbitrary number of input voters to return a positive

answer. Finally, 0.15 X of cT bundle allows us to compute a majority voting algorithm, set a clear discrimination between 4 and 6 inputs.



Extended Data Fig. 7 Operation of the NOR gate. a, A NOR gate is formed by hybridizing a template strand (accepting α or β as primers) with a molecular beacon γ. In the absence of α and β, the fluorescence of the molecular beacon is high because its dye and quencher are far apart. This codes for the ON state of the gate. If α or β is present, it induces polymerase-mediated displacement of the molecular beacon from the template, resulting in a low level of fluorescence. **b,** Verification of the assembly of the beacon with the template. The fluorescence of the beacon is monitored as the solution is cooled in a thermocycler, with and without template. The presence of the template induces a large fluorescence increase, due to the hybridization of the beacon. **c,** A two-layer network with the NOR gate as its output (left) and its truth table (middle) The network consists of two linear classifiers (α and β), each accepting a distinct input (X_1 or X_2). The correct operation of the network is verified by fluorescence: presence of either input drives the polymerization of the gate and the displacement of the beacon, leading to a drop in fluorescence.



Extended Data Fig. 8 Thermal dependence of the space partitioning network (Figure 5). Droplets with the network and varying inputs were incubated in a graded temperature field to reveal how the quality of classification varies with temperature. At low temperature (~43 °C), α and β coexist, as indicated by the yellow region. When temperature increases, the yellow region shrinks and disappears, indicating that α and β do not coexist.