nature nanotechnology

ARTICLES

PUBLISHED ONLINE: 29 SEPTEMBER 2013 | DOI: 10.1038/NNANO.2013.189

Programmable chemical controllers made from DNA

Yuan-Jyue Chen[1], Neil Dalchau[2], Niranjan Srinivas[3], Andrew Phillips[2], Luca Cardelli[2], David Soloveichik[4]* and Georg Seelig[1,5]*

Biological organisms use complex molecular networks to navigate their environment and regulate their internal state. The development of synthetic systems with similar capabilities could lead to applications such as smart therapeutics or fabrication methods based on self-organization. To achieve this, molecular control circuits need to be engineered to perform integrated sensing, computation and actuation. Here we report a DNA-based technology for implementing the computational core of such controllers. We use the formalism of chemical reaction networks as a 'programming language' and our DNA architecture can, in principle, implement any behaviour that can be mathematically expressed as such. Unlike logic circuits, our formulation naturally allows complex signal processing of intrinsically analogue biological and chemical inputs. Controller components can be derived from biologically synthesized (plasmid) DNA, which reduces errors associated with chemically synthesized DNA. We implement several building-block reaction types and then combine them into a network that realizes, at the molecular level, an algorithm used in distributed control systems for achieving consensus between multiple agents.

Molecular devices have captured the imagination of chemists and engineers for at least 30 years[1]. Rationally designed 'active' molecules include nanoparticles for the targeted delivery of drugs and imaging agents[2], or molecular motors that move along tracks and deliver cargo[3]. DNA nanotechnology[4,5] is in a unique position among the many actively pursued strategies for constructing molecular nanorobots, demonstrating progress towards the rational design of all the required elements: sensors and amplifiers[6–11], circuits[12–25], motors[26–30] and structures[4,31,32]. A rationally designed molecular robot has even combined structural elements with sensing and actuation, although it lacked complex embedded control[33]. The DNA-only construction of digital logic circuits and Boolean neural networks with over a hundred rationally designed parts forms possibly the most dramatic demonstration of a systematic engineering approach to building molecular circuits[16,17]. However, these approaches to constructing molecular information-processing systems do not realize the full spectrum of analogue and temporal dynamics naturally present in chemistry, which can be harnessed to control active molecular devices.

We experimentally demonstrate a design strategy for building DNA-only chemical controllers capable of being programmed to execute analogue temporal dynamics. The technology is designed around a signalling protocol based on short single-stranded DNA sequences. Molecular sensors (for example, aptamer switches) can release or expose such short sequences, and actuators (for example, antisense drugs or ribozymes) can be triggered by them. MicroRNAs can also be used as inputs to DNA circuits[18,34]. The control system we design sits in between, receiving inputs in the form of DNA sequences, and producing outputs in the form of other sequences (Fig. 1a). The treatment of controller, sensor and actuator as independent modules has proved indispensible in other fields of engineering.

Our DNA components are, in principle, capable of realizing the entire diversity of dynamic behaviours of chemical kinetics as mathematically captured by a chemical reaction network (CRN)[12,19]. Although CRNs started out as a tool to understand experimental observations of elementary chemical reactions, they form a general framework for modelling systems with many interacting components, such as gene regulatory networks, animal populations and sensor networks. CRNs can embody a wide range of digital and analogue behaviours, including temporal pattern generation, multistability and memory, Boolean logic, signal processing, control systems or distributed algorithms[13,35–40]. Moreover, viewed as a programming language, CRNs provide a natural and intuitive formalism for delineating and reasoning about molecular interactions, without making underlying physical details explicit.

We use the familiar language of chemistry to write programs for our DNA architecture (Fig. 1a). The 'instruction' $A + B \rightarrow C + D$ means that the signals A and B are transformed into signals C and D, where A, B, C and D are DNA strands we design. The reaction is not elementary; rather, it is systematically 'compiled' into a sequence of DNA strand displacement reactions. Our use of this chemical programming language is not gratuitous—a central contribution of this Article is to provide experimental evidence that our DNA architecture produces the expected stoichiometry and mass action kinetics of chemical reactions, so that our algorithms can behave similarly to what one might naively expect.

We test the major reaction classes—non-catalytic, catalytic and autocatalytic reactions. We then combine multiple such building blocks into a network implementing a distributed control algorithm for achieving consensus between multiple agents. Although the connection between distributed computing and chemistry has been noted many times in the literature (for example, Petri nets[41]), the sophistication of the molecular engineering required has deterred

[1]University of Washington Department of Electrical Engineering, 185 Stevens Way, Paul Allen Center – Room AE100R, Campus Box 352500, Seattle, Washington 98195-2500, USA, [2]Microsoft Research, 21 Station Road, Cambridge CB1 2FB, UK, [3]Computation and Neural Systems, California Institute of Technology, 1200 E California Boulevard, Mail Code 136-93, Pasadena, California 91125, USA, [4]Center for Systems and Synthetic Biology, University of California, 1700 4th Street, Byers Hall 401, Box 2540, San Francisco, California 94158, USA, [5]Department of Computer Science and Engineering, University of Washington, Box 352350, Seattle, Washington 98195-2350, USA. *e-mail: david.soloveichik@ucsf.edu; gseelig@uw.edu

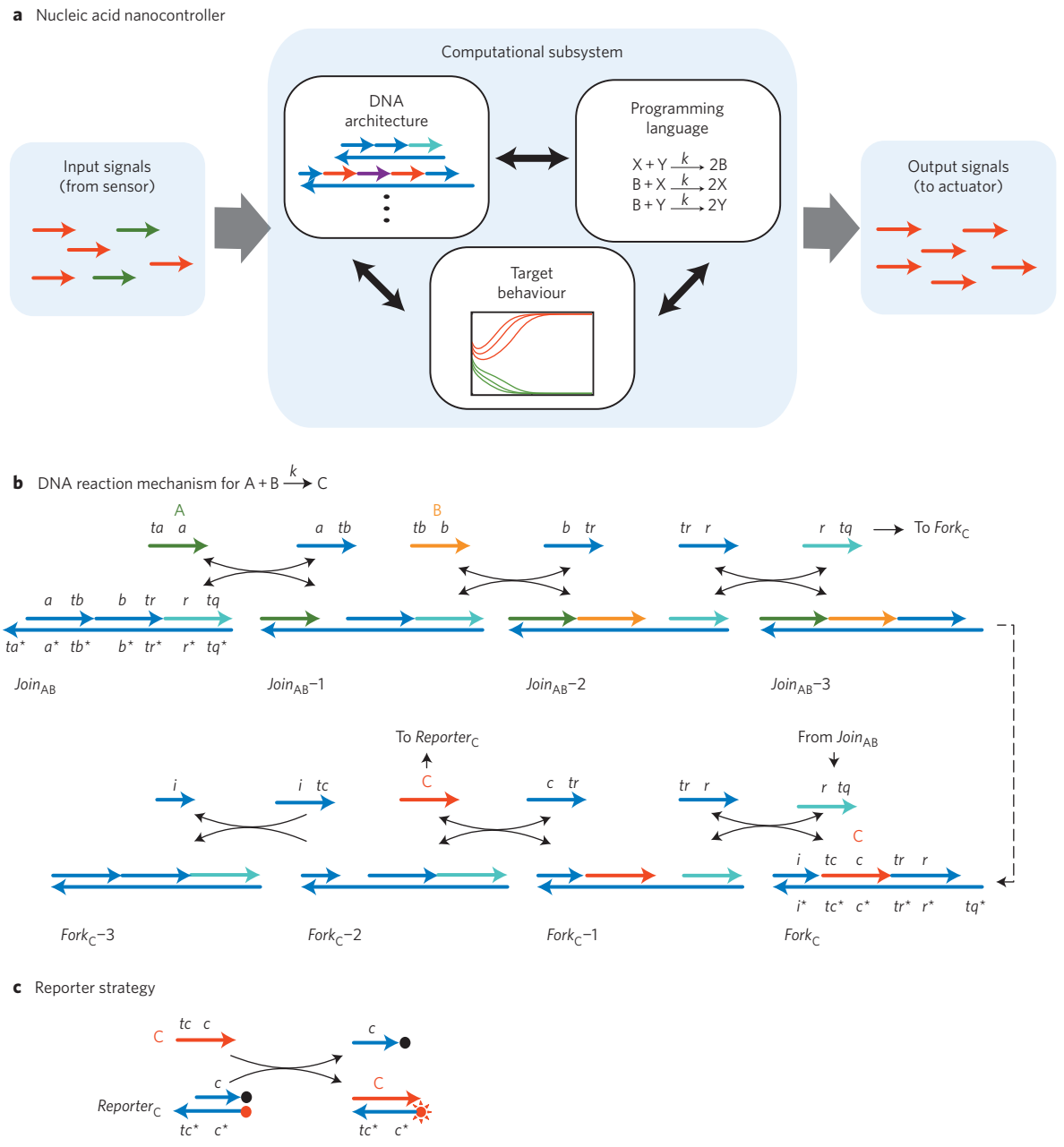


Figure 1 | DNA realization of a formal CRN. a, A standardized signalling protocol based on short single strands of DNA enables the components of the nanocontroller to communicate with each other. The formalism of CRNs serves as a programming language that specifies the desired behaviour for the computational subsystem. The target behaviour is experimentally realized by the DNA architecture. **b**, Reaction mechanism. DNA strands are drawn as lines with arrows at the 3′ end. Functional domains are labelled with lowercase letters; * indicates Watson–Crick complement. Species A, B and C of the formal reaction are represented by DNA signal strands *A* (⟨*ta a*⟩, green), *B* (⟨*tb b*⟩, orange) and *C* (⟨*tc c*⟩, red), respectively. Implementation of the bimolecular reaction A + B → C requires two multistranded gate complexes $Join_{AB}$ and $Fork_C$, as well as the auxiliary strands ⟨*tr r*⟩, ⟨*c tr*⟩ and ⟨*i tc*⟩. The reaction proceeds through a sequence of six strand displacement reactions, where each step provides a toehold for initiation of the next. **c**, Reporting strategy for reaction kinetics used in this Article. The reporter consists of two strands, one labelled with fluorophore (red dot) and the other with a quencher (black dot). Fluorescence is quenched when fluorophore and quencher are co-localized. Displacement of the quencher-labelled strand by signal C leads to an increase in fluorescence proportional to the amount of C detected.

experimental implementations. Our experiments corroborate that we can realize complex behaviours previously out of reach of synthetic molecular systems.

Among the many proposed architectures for strand displacement computation[2,10–13,15–19], ours is unique in that it relies exclusively on linear, double-stranded DNA complexes (processed by 'nicking' one of the strands)[10]. Because this structure is compatible with natural DNA, we are able to produce our computational elements in a highly pure form by bacterial cloning. Thus, we bypass the practical limitations in the length and purity of synthetic strands.

Signal transduction mechanism

We identify signals (A, B, C, ...) with single-stranded DNA molecules (signal strands, Fig. 1b). Nicked double-stranded DNA (ndsDNA) gate complexes mediate interactions between these signal strands with the help of additional auxiliary single-stranded

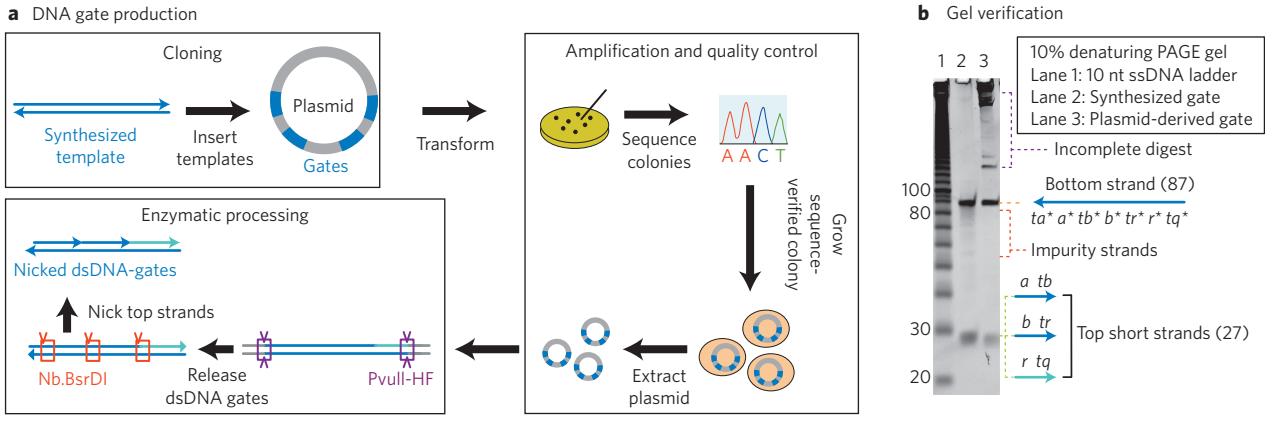


Figure 2 | DNA gate production. a, Highly pure ndsDNA gates can be produced from plasmid DNA. Multiple copies of a double-stranded ndsDNA gate template are inserted into a plasmid and transformed into *E. coli* cells. Clones are picked and plasmid sequence is verified. A clonal population is grown up, and plasmid DNA is extracted using standard molecular techniques. Finally, the restriction enzyme PvuII-HF is used to release the gate from the plasmid, and the nicking enzyme Nb.BsrDI is used to generate nicks in the top strand. **b**, Analysis by 10% denaturing polyacrylamide gel electrophoresis (PAGE) of the enzymatically processed gate. The long bottom strand (87-mer) and short top strands (27-mer) are visible on the D-gel.

species. All signal strands have the same sequence domain structure (see for example, signal strands A ($\langle ta\ a \rangle$, green), B ($\langle tb\ b \rangle$, orange) and C ($\langle tc\ c \rangle$, red) in Fig. 1b) with a short toehold domain (labels *ta*, *tb*, ...) that initiates binding to a gate, followed by a long domain (*a*, *b*, ...) that determines signal identity.

The reaction A + B → C is implemented with two gates (called $Join_{AB}$ and $Fork_C$ in Fig. 1b). The join gate consumes (and thus 'joins') the two signals A and B and the fork gate releases the signal C, which is initially bound to the fork gate $Fork_C$, and thus inactive. (The name 'fork gate' derives from the fact that multiple signal strands can be released, as shown in later examples.) The complete triggering of a join and a fork gate—corresponding to a single formal reaction A + B → C—is a cascade of strand displacement reactions in which each reaction exposes a toehold for the subsequent reaction (Fig. 1b, Supplementary Section S1). The displacing strand is a signal strand, an auxiliary strand or a strand previously released in the cascade (for example, 'translator' strand $\langle r\ tq \rangle$ is released by the join gate and triggers the fork gate). We use a fluorescent reporter strategy to detect specific strands and follow the reactions (Fig. 1c).

Each reaction is reversible until the very last displacement step involving the fork gate. The reversibility of the first step is essential to ensure stoichiometric correctness: the first formal reactant A should not be consumed in the absence of the second, B. Reversibility allows A to be re-released if the cascade does not complete.

The two-gate design and use of auxiliary strands ensures that all signal strands have the same domain structure and independent sequence, which guarantees composability[12]. Signal strands can thus be shared between multiple reactions to create a coupled system. Without these constraints, we can implement an individual bimolecular reaction with many fewer species, but we would lose the ability to compose reactions into arbitrary CRNs.

Plasmid encoding of DNA gates

The performance of strand displacement systems is currently limited by undesirable side reactions: leaks (the spontaneous 'firing' of a reaction cascade in the absence of the intended molecular trigger) or substoichiometric completion levels (unintentional sequestration of the signal leading to reduced product yield). These problems can at least in part be traced to errors in chemical DNA synthesis[42]. Biologically synthesized DNA is a useful alternative to synthetic DNA, even in non-biological applications where large quantities of highly pure DNA are required[43–45].

Our gates consist entirely of nicked double-stranded DNA[12], which makes them uniquely compatible with plasmid DNA as a starting material. Plasmid-derived gates have the additional advantage that they can be replicated and stored as bacterial glycerol stocks (before enzymatic processing). Gate production is detailed in Fig. 2a. Correct processing was tested using gel electrophoresis (Fig. 2b, Supplementary Section S7). Enzyme selection and additional design criteria are detailed in Supplementary Sections S2 and S3. The sequence constraints imposed by the use of nicking enzymes do not limit the generality of our method. Signals can be made orthogonal to one another by designing the sequences surrounding the nicking sites to be different. All data shown in this Article were collected with plasmid-derived ndsDNA gates except where otherwise indicated. Externally added signal and auxiliary strands, as well as the reporter gates used for following reaction kinetics, were chemically synthesized.

Testing fundamental reaction types

The modular nature of our design makes it easy to create reactions with multiple products of unconstrained sequence, allowing us to engineer the three major reaction classes: non-catalytic, catalytic and autocatalytic. These are the building blocks for composition of complex CRNs.

Extensive tests of the most basic reaction A + B → C verified correct stoichiometry (are the correct amounts of reactants used up and products generated?; Fig. 3a) and kinetics (are the reactants and products being generated according to the target rate law?; see section 'Verification of the bimolecular rate law'). In the catalytic reaction A + B → C + B, even a small amount of B effectively 'converts' all of A to C, but B remains conserved (Fig. 3b). Catalytic reactions are ubiquitous in biological chemical controllers (for example, transcriptional networks, kinase networks) as well as man-made artificial systems[6–11]. In Supplementary Fig. S10, we quantitatively analyse the catalytic turnover, showing that a single catalyst can trigger multiple reaction cycles.

In the autocatalytic reaction A + B → C + 2B, even a small amount of B effectively 'converts' all of A to itself (C acts as a 'readout'), resulting in the typical sigmoidal kinetic curves (Fig. 3c). Because of the exponential growth kinetics, autocatalytic reactions are common in settings where rapid (self-)amplification is observed, such as replication or apoptosis. These properties also make autocatalysis a key ingredient for propagating information in proposed chemical algorithms[46] (see also section 'Consensus network'). Because autocatalysis is extremely sensitive to leaks[9–11],

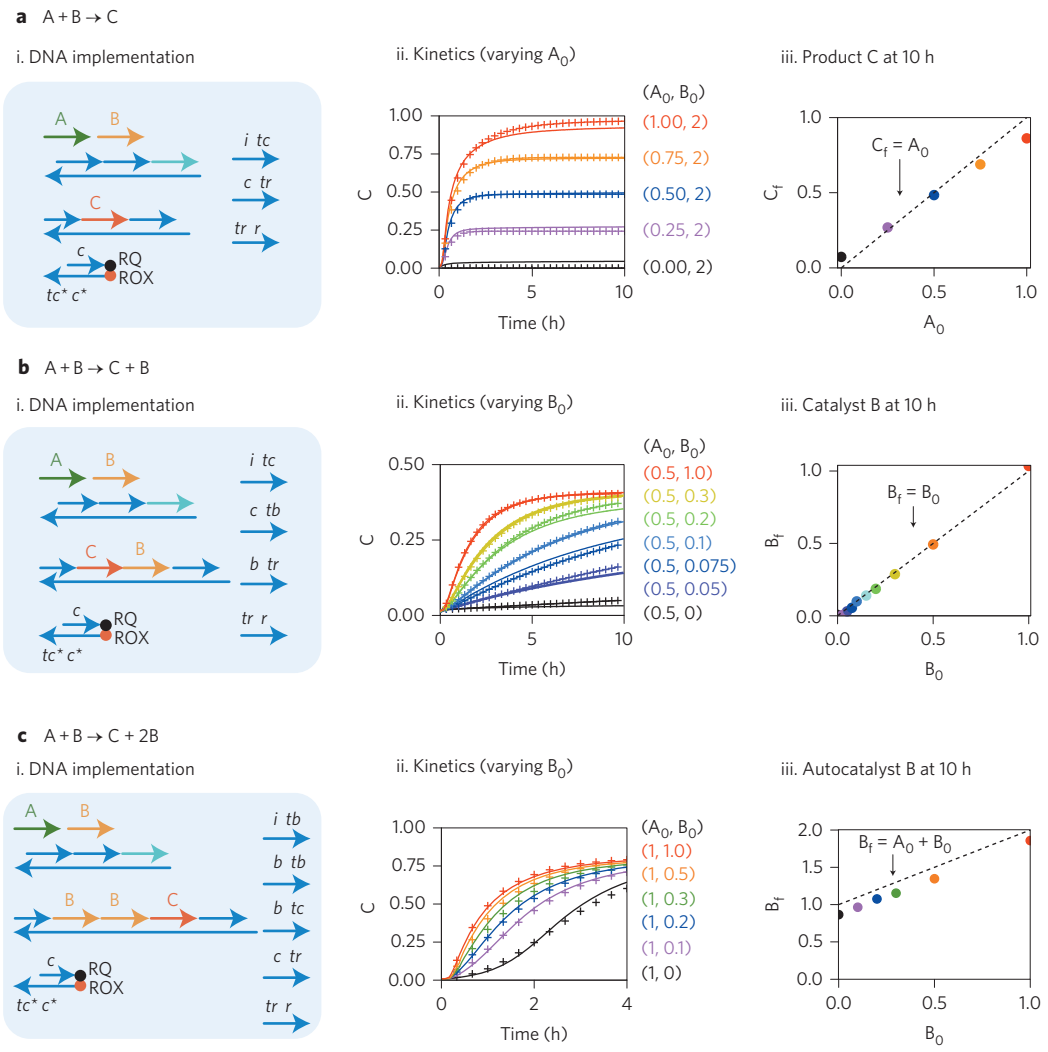


Figure 3 | Testing fundamental reaction types. Panels (i) show a simplified representation of the gates, auxiliary strands, and signal strands used for the corresponding experiments. Experimental kinetics data are shown in panels (ii) as full coloured lines. Concentrations of the signal strands are indicated in the same colour, 1× = 50 nM. All join and fork gates were at 1.5× , and auxiliary strands were at 2×. Best fits of the strand displacement-level model to the data are shown as crossed lines. Panels (iii) show data confirming the correct reaction stoichiometry. **a**, Non-catalytic bimolecular reaction A + B → C. Signal B was at 2× and different amounts of signal strand A were added. Panel (iii) shows that levels of (product) signal C at the measurement end point (10 h) are very close to the amounts of limiting inputs as expected for a stoichiometrically correct bimolecular reaction. **b**, Bimolecular catalytic reaction A + B → C + B. Signal A was at 0.5× and different amounts of the catalytic signal B were introduced into the system. Panel (iii) shows that the final amount of free catalyst B_f is equal to the initial amount B_0. The amount of catalyst signal B at 10 h was measured by adding a fluorescent reporter for B. **c**, Autocatalytic reaction A + B → C + 2B. Signal A was at 1× and the amount of signal B was varied. Panel (iii) shows that the final amount of the autocatalyst signal B is equal to the sum of the initial amounts of A and B as expected for autocatalysis. The amount of autocatalyst signal B was measured at 10 h by adding a fluorescent reporter for B.

it provides a good measurement of reactant quality. The estimated amount of autocatalyst B leaked (black trace, Fig. 3c,ii) is less than 2% (Supplementary Table S3); however, this leak is exponentially amplified.

To compare the performance of plasmid-derived gates to that of synthesized gates, we re-implemented the catalytic and autocatalytic reactions with synthesized gates using the same sequences. We observed that catalytic turnover is at least twice as high for the plasmid-derived gates. This observation is indicative of incomplete triggering due to unknown side reactions sequestering the catalyst in the synthesized system. Comparing autocatalytic reactions, we found that plasmid-derived gates suffered from noticeably less of the untriggered amplification characteristic of a lower leak rate (Supplementary Fig. S10). These data are consistent with the observation that there are fewer truncated strands detected in a gel analysis of the plasmid-derived gates than for the synthetic gates (Fig. 2b, Supplementary Section S3.3).

We tested bimolecular reactions with one, two or three products, but our approach can be generalized to different numbers of products and reactants. Unimolecular reactions can be implemented with a single-input join gate, while higher-order reactions can be implemented using join gates with multiple inputs.

Verification of the bimolecular rate law

The reaction specification A + B → C delineates not only the production/consumption relationships between A, B and C, but also the dynamics. Despite the overall complex reaction mechanism (which, for A + B → C, involves five reversible and one irreversible stand displacement reactions, Fig. 1b), an analytical argument shows that the overall kinetics should be well approximated by the

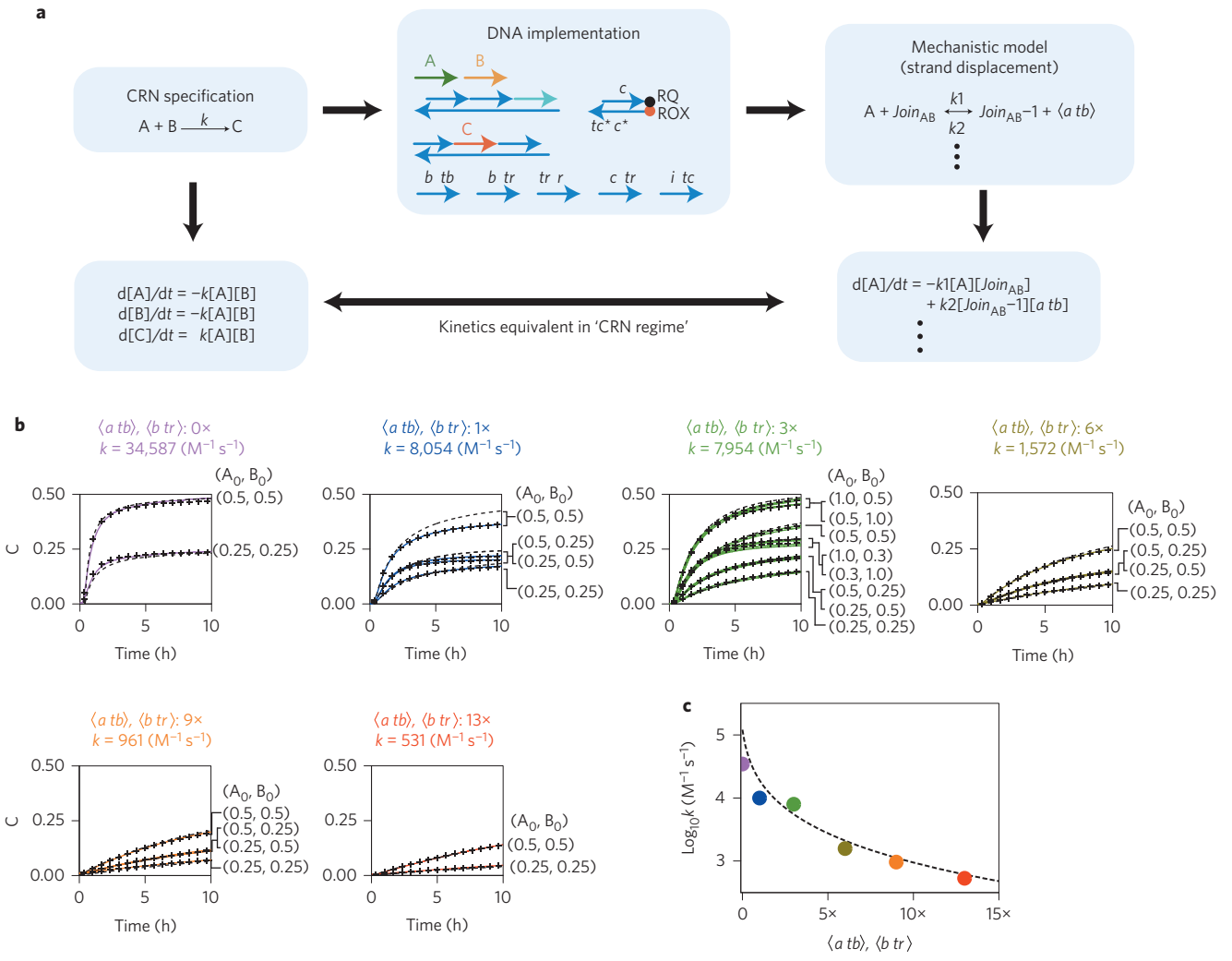


Figure 4 | Tuning the rate of the bimolecular reaction A + B → C. a, Approximating the bimolecular rate law. The CRN program is executed by a DNA architecture that can be quantitatively modelled at a mechanistic level. We view strand displacement reactions (for example, A + $Join_{AB}$ → $Join_{AB}$−1 + ⟨*a tb*⟩, see Fig. 1b for component names) as the *elementary* reaction steps and the formal bimolecular reaction (for example, A + B → C) as the *complex* reaction pathway decomposed into these elementary reactions. In the 'CRN regime' (see text) the mechanistic model closely approximates the dynamics of the target program. The rate constant *k* of the formal system can be tuned by changing the concentrations of gates and auxiliary strands. **b**, Reactions with varying concentrations of the backward auxiliary strands ⟨*a tb*⟩ and ⟨*b tr*⟩. The data (solid traces) show the time evolution of C; purple traces (0× ⟨*a tb*⟩ and ⟨*b tr*⟩), blue (1×), green (3×), olive green (6×), orange (9×) and red (13×), where 1× = 40 nM. Gates were at 3× and the initial concentrations of signals A and B are indicated in each panel. Black dashed lines are fits to the bimolecular rate law in **a**. Best-fit rate constants to the bimolecular rate law are indicated in each panel. Black crossed lines are fits to the mechanistic strand displacement-level model. **c**, Fitted bimolecular rate constant versus analytic prediction. The solid line is obtained from an analytic prediction for the dependence of the expected rate constant on the concentrations of the backward auxiliary strands ⟨*a tb*⟩ and ⟨*b tr*⟩ (equation (8) in Supplementary Section S5). The coloured dots show the rate constants *k* obtained from fitting the experimental data from **b**.

mass-action rate law expected of the formal reaction (that is, $d[C]/dt = -d[A]/dt = -d[B]/dt = k[A][B]$). As the derivation in Supplementary Section S5 shows, the regime of best correspondence ('CRN regime') is one in which gates and auxiliary strands, including 'backward' auxiliary strands ⟨*a tb*⟩ and ⟨*b tr*⟩, are sufficiently in excess over the signal strands (Fig. 4a).

We experimentally confirmed that the multistep strand displacement level mechanism implements the expected rate law for A + B → C, and that the rate constant can be tuned by adjusting the concentrations of gates and auxiliary species. Figure 4b shows six sets of experimental data for the reaction A + B → C in or near the CRN regime. Each set was obtained with a different concentration of the backward auxiliary strands ⟨*a tb*⟩ and ⟨*b tr*⟩ and contains kinetics traces corresponding to at least two different combinations of the signals A and B. We chose to vary the concentration of the backward auxiliary strands because our analysis suggests that the formal rate constant can be effectively tuned in this way (Supplementary Sections S5). We then fit the data from each set to a bimolecular rate law. The best fit rate constants varied over about two orders of magnitude from $3.5 \times 10^4\ M^{-1}\ s^{-1}$ to $5.3 \times 10^2\ M^{-1}\ s^{-1}$ as the concentration of backward auxiliary strands increased from 0× to 13× (Supplementary Section S5). The data show that the reactions are symmetrical with regard to the two signals, as required by the bimolecular rate law, although signal strands A and B react sequentially with the join gate (see, for example, traces with A, B at 1×, 0.3× and 0.3×, 1× respectively).

Mechanistic strand displacement-level model

Each individual strand displacement step can be well modelled as a bimolecular reaction between a signal or auxiliary strand and a gate complex with a matching open toehold[47]. We used the Visual DSD[14,48] software to develop a quantitatively constrained model of

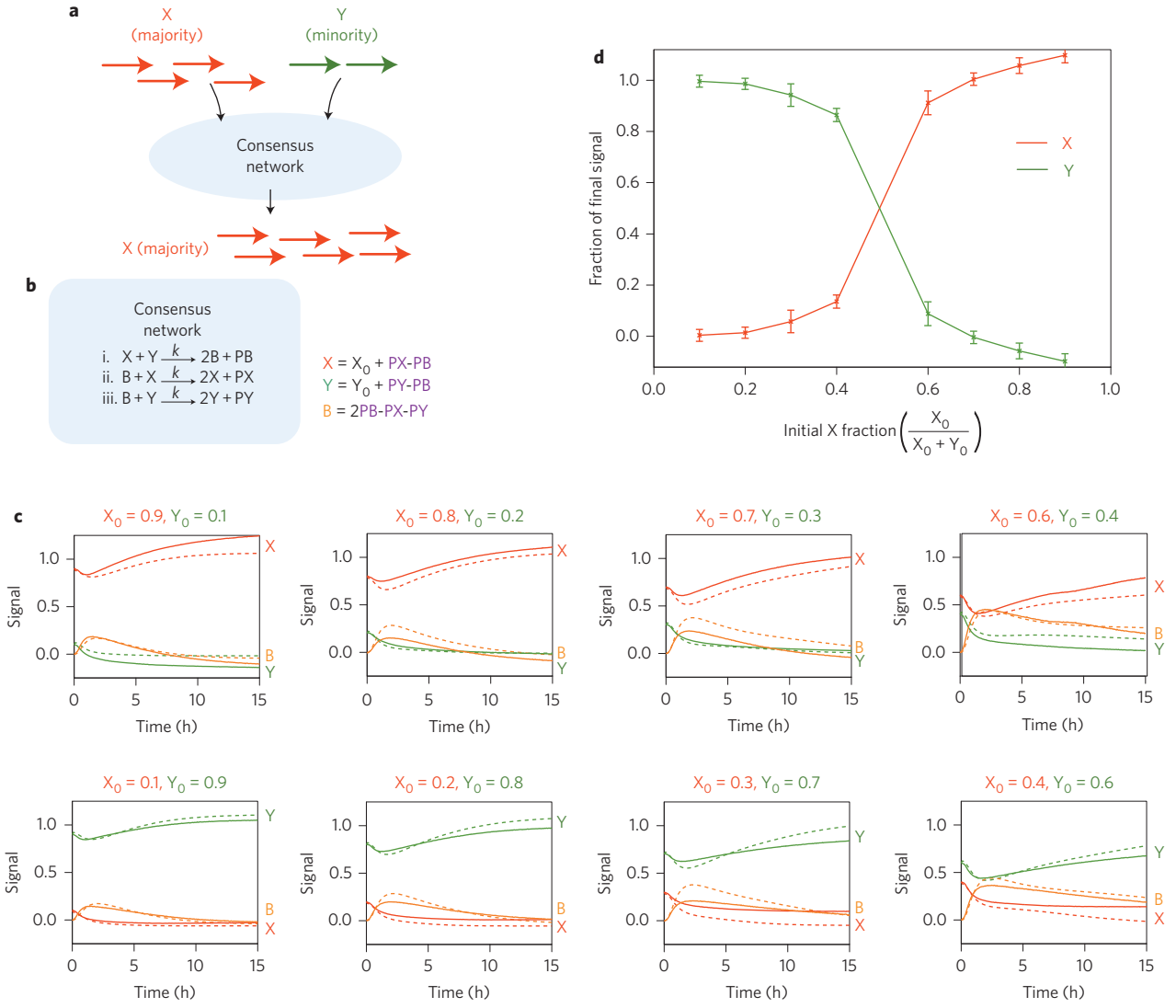


Figure 5 | Consensus network. a, Given arbitrary amounts of signal strands X (red) and Y (green), the consensus network converts the minority signal to the majority signal. **b**, The formal chemical reactions for the consensus network. Signals PX, PY and PB were used to follow the reaction kinetics without interfering with the dynamics of X, Y and B. Reporters for PX, PY and PB each used a different fluorophore such that all three signals could be detected in the same reaction. The values of X, Y and B were calculated from the measured values of PX, PY and PB as indicated. **c**, Time evolution of signals X (red), Y (green) and B (yellow). Initial concentrations of signals X and Y are indicated in each panel, 1× = 80 nM. Reporters were at 3× , auxiliary strands at 2× and gates at 2× for reactions (i) and (ii). Gates and auxiliary strands for reaction (iii), B + Y → 2Y, were at 2.4× to balance the rates of the two autocatalytic reactions. The DNA implementation for the consensus network consisted of 3 join gates, 3 fork gates, 3 reporters, 13 auxiliary strands and 3 signal strands. No backward auxiliary strands were added to the initial reaction mixture. A graphical representation of all gates and auxiliary species is given in Supplementary Section S8.2. The kinetics data show that the minority species was converted into the buffer species B first, then into the majority species. The model prediction of the consensus network using the strand displacement-level model is shown as dashed lines. The prediction is based on a model parameterization obtained by fitting to the individual reactions (Supplementary Section S8). **d**, Amplification levels. The end points (15 h) of each reaction show that the DNA-based consensus network correctly amplifies the majority towards totality. Red trace: X/(X + Y) at 15 h; green trace: Y/(X + Y) at 15 h. Error bars indicate standard deviation calculated from three independent experiments.

the dynamics of our system on this mechanistic level. We allowed each strand displacement step to occur at a different rate depending on the sequences of the toeholds and adjacent domains. The model includes a phenomenological treatment of the erroneous leak reactions (Supplementary Section S7.3). We fit all the data that we obtained for the non-catalytic, catalytic and autocatalytic reactions (Figs 3 and 4) and independent measurements of a large number of intermediate reaction steps (Supplementary Figs S17,S18). These 104 data traces yielded a highly constrained set of strand displacement rate constants, with values ranging from $1\times10^{4}\ M^{-1}\ s^{-1}$ to $1.44\times10^{6}\ M^{-1}\ s^{-1}$ (Supplementary Table S3), consistent with previously reported data[47]. The mechanistic strand displacement-level model fitted the data for all reaction conditions, including leak reactions, exceptionally well (crossed lines in Figs 3 and 4b, Supplementary Section S7).

The strand displacement rate constants can be used to predict the effective bimolecular rate constant for the target formal reaction A + B → C (Supplementary Section S5). We compared this predicted rate constant to that obtained by direct fitting of the data in Fig. 4b to a bimolecular rate law. Figure 4c shows that our prediction is in good agreement as long as the concentration of the backward strands is 3× or higher. Divergence is expected at lower concentrations because the approximation we made to derive the analytic result does not hold. These results demonstrate

that we can systematically vary the formal rate constant through quantitative control over the underlying reaction mechanism.

Consensus network

An important function of molecular controllers is their ability to make decisions by comparing concentrations of input signals (Fig. 5a). We engineered such a decision-making controller by implementing a consensus algorithm that operates on two signals (X and Y). The signal that is initially in the minority is completely eliminated and replaced by the signal that is initially in the majority (in any given experiment the sum of all signal concentrations is constant) (Supplementary Section S8.1). In distributed computing parlance, we implement an algorithm that allows picomole quantities of agents, each with vote X or Y, to agree on a majority decision[46,49]. The classification into minority and majority is thus unequivocal, distinguishing this network from previous proposals for DNA-based molecular classifiers[50] where the resulting signal was proportional to the difference in the initial concentrations (which can be small when concentrations are close). Consensus is a basic distributed computing problem and provides for us a proof-of-principle that CRN algorithms are directly translatable to our DNA controllers.

The network consists of two autocatalytic and one non-catalytic bimolecular reaction (Fig. 5b): a reaction between signals X and Y creates two copies of the buffer signal B, while a reaction involving B and X(Y) creates two copies of X(Y). Intuitively, the minority and majority signals initially cancel each other, producing the buffer signal, which is then converted back to the majority signal. For the threshold to be set at the point where concentrations of the two species are equal, the two autocatalytic reactions should occur at the same rate. To compensate for minor differences in the rates of the two autocatalytic reactions (Supplementary Section 8.2) and make reaction rates similar, we adjusted gate and auxiliary species concentrations.

Supplementary Fig. S21 shows the ndsDNA gates needed to implement this reaction network (sequences and other design considerations are listed in Supplementary Section S2 and Table S6). Example data traces in Fig. 5c clearly show that the network correctly classifies the majority for all eight input combinations tested. Each panel shows the time evolution of the concentration of X (red), Y (green) and B (yellow) for given starting concentrations of X and Y. Figure 5d shows the net amplification of the majority signal relative to the minority signal. The results for the DNA implementation and the expected dynamics of the formal CRN (Supplementary Fig. S20) are in qualitative agreement: we observe the gradual decrease of the minority, intermediate build-up of buffer and the initial decrease but long-term increase of the majority.

We also constructed a strand displacement model for each reaction of the consensus network using Visual DSD, and parameterized these models using experimental measurements for each reaction and for the individual fork and join gates (Supplementary Figs S21, S23, S24). By composing models of individual reactions into a model of the full consensus network, we were able to quantitatively predict the dynamics of the consensus network solely from the models of its constituent parts, up to a constant scaling factor (Fig. 5c; see Supplementary Section S8 for further details).

Conclusions

We have developed a new systematic design strategy for non-living molecular systems with functional behaviours, paired with a technology for robust and efficient synthesis of the molecular components. Our scheme is built upon *de novo* designed interactions not known to occur in nature. Our components did not require (directed) evolution to achieve efficacy, but were designed in their ultimate form by the authors. As such, our work can be seen as a step in the larger human enterprise of recapitulating the mastery of biology over matter with *de novo* engineering.

As human engineering is driven inexorably towards molecular-scale devices, we must be careful to avoid shoehorning theory developed for digital electronics (for example, logic circuits) into the chemical context. For well over a century, CRNs have provided the mathematical language to describe and predict the dynamics of chemical experiments. Here, we leveraged this significant theoretical investment and demonstrated the prescriptive use of CRNs for programming molecular nanocontrollers.

Although our devices are entirely synthetic, they are biocompatible and there is a natural path towards applications in sensing and smart drug-delivery *in vivo*. Cell state is encoded in the sequences and concentrations of RNAs, and recent work has shown that strand displacement logic gates can recognize miRNA profiles in living mammalian cells[34]. Further, the demonstration of the power of strand displacement as a mechanism for building synthetic molecular circuits tempts the hypothesis that there are natural strand displacement-based cellular regulatory networks with interesting dynamics yet to be discovered.

Methods

Preparation of plasmid-derived ndsDNA gates. Double-stranded DNA templates were cloned into a high-copy-number plasmid and transformed into *Escherichia coli*. A single colony was picked from an Ampicillin selective plate and an 800 ml overnight culture was grown with Ampicillin (100 μg ml^{-1}) at 37 °C with vigorous shaking. Plasmids were extracted using a QIAGEN Maxi-prep kit, and inserts were sequenced to ensure that there was no sequence error or recombination in the ndsDNA gates. Cloned ndsDNA gates were first digested with a restriction enzyme (PvuII-HF) at 37 °C for 1 h to release the gates from the plasmid backbone. Reactions were run with 4 units of PvuII-HF per 1 μg of plasmid. The reaction mix was then ethanol precipitated to optimize the reaction conditions for the next digestion step. Join gates were digested with the nicking enzyme Nb.BsrDI at 65 °C for 1 h using 4 units of enzyme per 1 μg of plasmid. Fork gates were digested with the nicking enzyme Nt.BstNBI at 55 °C for 1 h using 8 units of enzyme per 1 μg of plasmid (for enzyme amount optimization see Supplementary Sections S3.2 and S3.4). All enzymes were purchased from New England Biolabs. For kinetics experiments, enzymes were dissociated from DNA by adding sodium dodecyl sulphate (SDS) to a final concentration of 0.15% (Supplementary Section S3.5). Gates were then used for experiments without further purification from enzymes or plasmid backbone.

Modelling and parameter inference. Computational models were constructed for each analysed circuit using the DNA Strand Displacement (DSD) programming language and Visual DSD software[14,48]. The unknown kinetic parameters in the model were inferred from the experimental data using Markov chain Monte Carlo methods, as implemented in the Filzbach software (see authors' website: http://research.microsoft.com/science/tools). Such methods require the definition of a likelihood function, which describes the probability of reproducing the observed data D, given a model hypothesis H and corresponding parameter set θ, that is, $\Pr(D|\theta, H)$. We used ordinary differential equation simulations for each circuit. C# code was generated using the Visual DSD tool, then integrated numerically with adaptive step-size ODE integrators (http://mstlab.org/eng/projects/Pages/Solvers.aspx Microsoft Research Solvers library for .NET). For further details, see Supplementary Section S7.

Further materials and methods can be found in Supplementary Section S9.

Received 5 March 2013; accepted 21 August 2013;
published online 29 September 2013

References

1. Drexler, K. E. Molecular engineering: an approach to the development of general capabilities for molecular manipulation. *Proc. Natl Acad. Sci. USA* **78,** 5275–5278 (1981).
2. Koo, O. M., Rubinstein, I. & Onyuksel, H. Role of nanotechnology in targeted drug delivery and imaging: a concise review. *Nanomedicine: NBM* **1,** 193–212 (2005).
3. Hess, H. Engineering applications of biomolecular motors. *Annu. Rev. Biomed. Eng.* **13,** 429–450 (2011).
4. Seeman, N. C. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79,** 65–87 (2010).
5. Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chem.* **3,** 103–113 (2011).
6. Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proc. Natl Acad. Sci. USA* **101,** 15275–15278 (2004).

7. Seelig, G., Yurke, B. & Winfree, E. Catalyzed relaxation of a metastable DNA fuel. *J. Am. Chem. Soc.* **128,** 12211–12220 (2006).
8. Turberfield, A. J. *et al.* DNA fuel for free-running nanomachines. *Phys. Rev. Lett.* **90,** 118102 (2003).
9. Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451,** 318–322 (2008).
10. Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318,** 1121–1125 (2007).
11. Levy, M. & Ellington, A. D. Exponential growth by cross-catalytic cleavage of deoxyribozymogens. *Proc. Natl Acad. Sci. USA* **100,** 6416–6421 (2003).
12. Cardelli, L. Two-domain DNA strand displacement. *Math. Struct. Comput. Sci.* **23,** 247–271 (2013).
13. Oishi, K. & Klavins, E. Biomolecular implementation of linear I/O systems. *IET Syst. Biol.* **5,** 252–260 (2011).
14. Phillips, A. & Cardelli, L. A programming language for composable DNA circuits. *J. R. Soc. Interface* **6,** S419–S436 (2009).
15. Qian, L., Soloveichik, D. & Winfree, E. Efficient Turing-universal computation with DNA polymers. *DNA Comput. Mol. Program.* **6518,** 123–140 (2011).
16. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332,** 1196–1201 (2011).
17. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475,** 368–372 (2011).
18. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314,** 1585–1588 (2006).
19. Soloveichik, D., Seelig, G. & Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc. Natl Acad. Sci. USA* **107,** 5393–5398 (2010).
20. Stojanovic, M. N. & Stefanovic, D. A deoxyribozyme-based molecular automaton. *Nature Biotechnol.* **21,** 1069–1074 (2003).
21. Benenson, Y. *et al.* Programmable and autonomous computing machine made of biomolecules. *Nature* **414,** 430–434 (2001).
22. Kim, J. & Winfree, E. Synthetic *in vitro* transcriptional oscillators. *Mol. Syst. Biol.* **7,** 465 (2011).
23. Montagne, K., Plasson, R., Sakai, Y., Fujii, T. & Rondelez, Y. Programming an *in vitro* DNA oscillator using a molecular networking strategy. *Mol. Syst. Biol.* **7,** 466 (2011).
24. Willner, I., Shlyahovsky, B., Zayats, M. & Willner, B. DNAzymes for sensing, nanobiotechnology and logic gate applications. *Chem. Soc. Rev.* **37,** 1153–1165 (2008).
25. Ran, T., Kaplan, S. & Shapiro, E. Molecular implementation of simple logic programs. *Nature Nanotech.* **4,** 642–648 (2009).
26. Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465,** 206–210 (2010).
27. Omabegho, T., Sha, R. & Seeman, N. C. A bipedal DNA Brownian motor with coordinated legs. *Science* **324,** 67–71 (2009).
28. Wickham, S. F. J. *et al.* Direct observation of stepwise movement of a synthetic molecular transporter. *Nature Nanotech.* **6,** 166–169 (2011).
29. Muscat, R. A., Bath, J. & Turberfield, A. J. A programmable molecular robot. *Nano Lett.* **11,** 982–987 (2011).
30. 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).
31. Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).
32. Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394,** 539–544 (1998).
33. Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335,** 831–834 (2012).
34. Hemphill, J. & Deiters, A. DNA Computation in mammalian cells: microRNA logic operations. *J. Am. Chem. Soc.* **135,** 10512–10518 (2013).
35. Arkin, A. & Ross, J. Computational functions in biochemical reaction networks. *Biophys. J.* **67,** 560–578 (1994).
36. Epstein, I. R. & Pojman, J. A. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford Univ. Press, 1998).
37. Magnasco, M. O. Chemical kinetics is Turing universal. *Phys. Rev. Lett.* **78,** 1190–1193 (1997).
38. Senum, P. & Riedel, M. Rate-independent constructs for chemical computation. *PLoS ONE* **6,** e21414 (2011).
39. Soloveichik, D., Cook, M., Winfree, E. & Bruck, J. Computation with finite stochastic chemical reaction networks. *Nat. Comput.* **7,** 615–633 (2008).
40. Tyson, J. J., Chen, K. C. & Novak, B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15,** 221–231 (2003).
41. Peterson, J. L. *Petri Net Theory and the Modeling of Systems* 290 (Prentice-Hall, 1981).
42. Zhang, D. Y. & Winfree, E. Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Res.* **38,** 4182–4197 (2010).
43. Lin, C. *et al. In vivo* cloning of artificial DNA nanostructures. *Proc. Natl Acad. Sci. USA* **105,** 17626–17631 (2008).
44. Ducani, C., Kaul, C., Moche, M., Shih, W. M. & Högberg, B. Enzymatic production of 'monoclonal stoichiometric' single-stranded DNA oligonucleotides. *Nature Methods* **10,** 647–652 (2013).
45. Chen, X., Briggs, N., McLain, J. R. & Ellington, A. D. Stacking nonenzymatic circuits for high signal gain. *Proc. Natl Acad. Sci. USA* **110,** 5386–5391 (2013).
46. Angluin, D., Aspnes, J. & Eisenstat, D. A simple population protocol for fast robust approximate majority. *Distrib. Comput.* **21,** 87–102 (2008).
47. Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* **131,** 17303–17314 (2009).
48. Lakin, M. R., Youssef, S., Cardelli, L. & Phillips, A. Abstractions for DNA circuit design. *J. R. Soc. Interface* **9,** 470–486 (2012).
49. Cardelli, L. & Csikász-Nagy, A. The cell cycle switch computes approximate majority. *Sci. Rep.* **2,** 656 (2012).
50. Zhang, D. & Seelig, G. in *DNA Computing and Molecular Programming* (eds Sakakibara, Y. & Mi, Y.) Vol. 6518, 176–186 (Lecture Notes in Computer Science, Springer, 2011).

Acknowledgements

The authors thank E. Winfree, E. Klavins and D.Y. Zhang for discussions and comments on the manuscript. This work was supported by the National Science Foundation (grant NSF-CCF 1117143 to G.S. and D.S.). G.S. was supported by a Burroughs Wellcome Career Award at the Scientific Interface. D.S. was supported by an NIGMS Systems Biology Center grant (P50 GM081879).

Author contributions

All authors designed the experiments and co-wrote the paper. Y-J.C. performed the wetlab experiments. N.D. and A.P. performed computational experiments.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.S. and G.S.

Competing financial interests

The authors declare no competing financial interests.

LETTER

doi:10.1038/nature12774

Experimental evidence for the influence of group size on cultural complexity

Maxime Derex[1], Marie-Pauline Beugin[1], Bernard Godelle[1] & Michel Raymond[1,2]

The remarkable ecological and demographic success of humanity is largely attributed to our capacity for cumulative culture[1–3]. The accumulation of beneficial cultural innovations across generations is puzzling because transmission events are generally imperfect, although there is large variance in fidelity. Events of perfect cultural transmission and innovations should be more frequent in a large population[4]. As a consequence, a large population size may be a prerequisite for the evolution of cultural complexity[4,5], although anthropological studies have produced mixed results[6–9] and empirical evidence is lacking[10]. Here we use a dual-task computer game to show that cultural evolution strongly depends on population size, as players in larger groups maintained higher cultural complexity. We found that when group size increases, cultural knowledge is less deteriorated, improvements to existing cultural traits are more frequent, and cultural trait diversity is maintained more often. Our results demonstrate how changes in group size can generate both adaptive cultural evolution and maladaptive losses of culturally acquired skills. As humans live in habitats for which they are ill-suited without specific cultural adaptations[11,12], it suggests that, in our evolutionary past, group-size reduction may have exposed human societies to significant risks, including societal collapse[13].

The accumulation of socially learned information over many generations has enabled humans to develop powerful technologies that no individual could have invented alone[14]. Cumulative culture is most likely to be restricted to the *Homo* genus and remains an evolutionary puzzle[15]. Several hypotheses have been proposed to explain this explosion in cultural complexity, with a recent emphasis on social-learning mechanisms specific to humans, such as teaching, language or imitation[16,17]. These mechanisms of faithful transmission stabilize cultural knowledge, thus enabling successive improvements, as has been previously shown theoretically[18] and empirically[19,20]. However, perfect transmission is most probably unrealistic, as for any given transmission event, an information loss is expected, particularly for complex tasks[4,21]. Moreover, transmission is only one aspect of the problem, as cumulative cultural evolution also requires the creation of new knowledge; that is, innovation.

The determinants of technological regression—the opposite situation—have been studied in Tasmanian aboriginals. It was argued that cultural losses were associated with population-size reduction[22]. A general model of cultural evolution that links demographic factors to psychological aspects of social learning has been proposed by Henrich[4]. Considering that transmission events for complex tasks are generally imperfect, with a large variance in fidelity, a learner could acquire by chance greater skill than the demonstrator if the number of transmission events (that is, the population size) is sufficiently large. As there is a psychological propensity to imitate successful individuals (prestige bias), this individual becomes the new demonstrator, driving cultural evolution. A decrease in population size makes such events unlikely, making cultural regression unavoidable. Analytical modelling shows that, as the population size increases, the combination of imperfect learning and prestige bias can lead to cumulative evolution, even if transmission is generally inaccurate. Bursts of cultural complexity during the Palaeolithic era (2.6 million years ago to 10 thousand years ago) and particularly during the Upper Palaeolithic transition (45 thousand years ago) may illustrate demographic processes, rather than changes in cognitive abilities[5,23]. However, factors favouring the ability to develop complex culture will most probably also have a positive effect on population size, thus limiting causal assessments using correlative studies. Furthermore, studies using anthropological data produced mixed results[6–9]. The only experimental study to investigate how group size influences cumulative cultural evolution reported no relationship[10]. However, only one cultural task was considered, and the larger group size was limited to three individuals. More parameters must be explored experimentally to investigate the effect of group size on cultural complexity.

Following Henrich's analysis, the maintenance of a cultural task within a group should depend on group size and task complexity. Specifically, within a group of a particular size, greater loss of information is expected for a more complex task. Alternatively, for a task of a given complexity, greater loss of information is expected in a smaller group. Thus, when considering two improvable tasks, one simple and one complex, artificially introduced into groups of different sizes, we predict that the simple task will be better conserved than the complex task (prediction 1); the probability of conserving the complex task will increase with group size (prediction 2); and better performance will be observed in the larger groups for both tasks (prediction 3).

To study the effect of group size on cultural complexity, 366 men participated in a dual-task computer game. Players had to collect resources individually to improve their 'health'. A cultural package composed of two demonstrations, one concerning a simple task and one concerning a complex task, was introduced within groups of different sizes (2, 4, 8 or 16 players). The players were told that each item in the cultural package could be improved. During each of the 15 trials of the game, each player had to build an arrowhead (simple task) or a fishing net (complex task) to collect 'life units' (see Extended Data Fig. 1). The cultural trait diversity of the group thus consisted of some players building one artefact, while the remaining players built the other; diversity was lost when all individuals built the same object.

As expected from prediction 1, the simple task was more likely than the complex task to be maintained for all group sizes ($\chi^2 = 3.83$, d.f. = 1, $P = 0.05$; Fig. 1a, b). For each task, the probability of being lost (none of the individuals of the group exploited the task at the end of the game, see Methods) by a group decreased with increasing group size ($\chi^2 = 7.62$, d.f. = 1, $P = 0.006$), as expected from prediction 2 (Fig. 1a, b). Interestingly, the increased probability of maintaining the complex task in large groups did not reduce the probability of maintaining the simple task (type of task × group size interaction $\chi^2 = 0.85$, d.f. = 1, $P = 0.36$). Indeed, the probability of maintaining cultural diversity (that is, observing both tasks in the group) increased with group size ($\chi^2 = 16.3$, d.f. = 1, $P < 0.0001$; Fig. 1c).

For each group size, the performances of the best within-group artefacts (simple and complex) at the fifteenth trial were compared to the score of the equivalent artefact from the cultural package. The simple task was stable in the smaller groups and improved in the larger groups (Fig. 2). A linear model was used to investigate the effect of

[1]University of Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France. [2]CNRS, Institute of Evolutionary Sciences, CC 065, Place Eugène Bataillon, Montpellier, France.

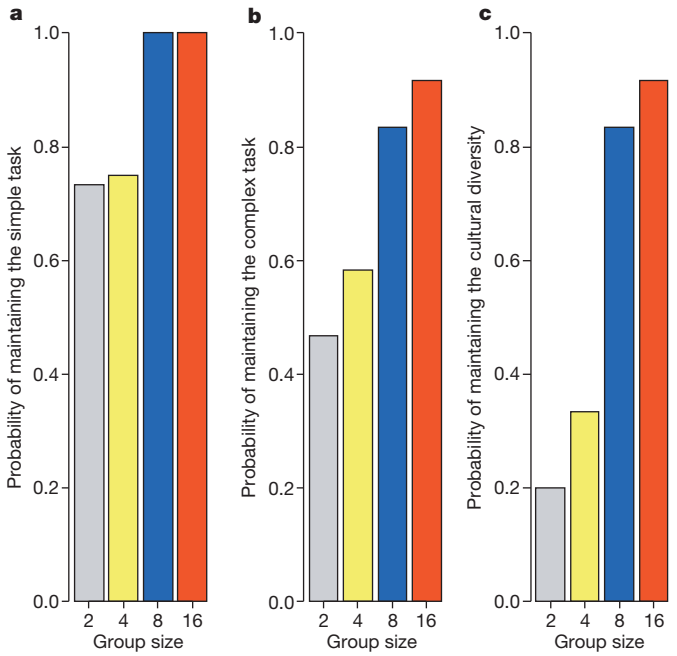


Figure 1 | Group size affects the maintenance of cultural tasks. **a**–**c**, Probability of at least one observation of the simple task (**a**), the complex task (**b**) or both (that is, cultural diversity) (**c**) among the three last trials, for group size of 2 ($n = 15$ replicates), 4 ($n = 12$), 8 ($n = 12$) and 16 ($n = 12$) players.

group size and shows that group size had a linear effect on the performance of the best within-group arrowhead, suggesting that cultural evolution was enhanced in larger groups, consistent with prediction 3 ($F_{1,48} = 10.2$, $P = 0.003$; Fig. 2 and Extended Data Fig. 2). Performance of the complex task deteriorated in the smaller groups and remained stable in the larger groups (Fig. 3). Group size had a linear and quadratic effect on the performance of the best within-group fishing net ($F_{1,47} = 7.12$, $P = 0.01$ and $F_{1,47} = 4.22$, $P = 0.05$, respectively; Fig. 3). Among groups maintaining the complex task, only the 8- and 16-player groups improved it compared to the original cultural package (see Extended Data Figs 3 and 4).

The improvement of both tasks was linked to group size, suggesting that refinement of pre-existing technology is facilitated by increasing group size. The link between innovation rate and group size is not

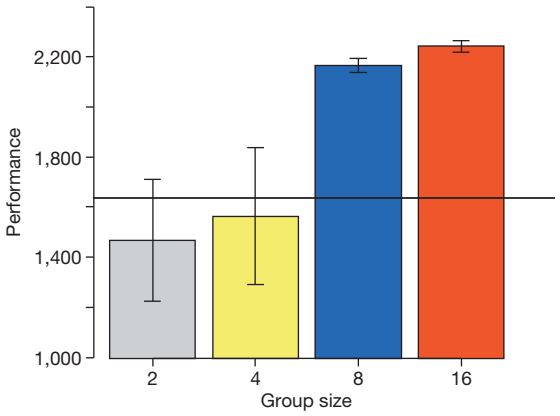


Figure 2 | Larger groups favour improvements to the simple cultural trait. The horizontal line shows the arrowhead performance from the cultural package. Performance is measured using arbitrary life units. Plotted are the mean values ± s.e.m. The simple task was stable in the smaller groups (mean performance: 2-player groups = 1,466, $t = -0.71$, d.f. = 14, $P = 0.49$; 4-player groups = 1,563, $t = -0.27$, d.f. = 11, $P = 0.79$) and improved in the larger groups (8-player groups = 2,166, $t = 18.84$, d.f. = 11, $P < 0.0001$; 16-player groups = 2,242, $t = 27.57$, d.f. = 11, $P < 0.0001$).

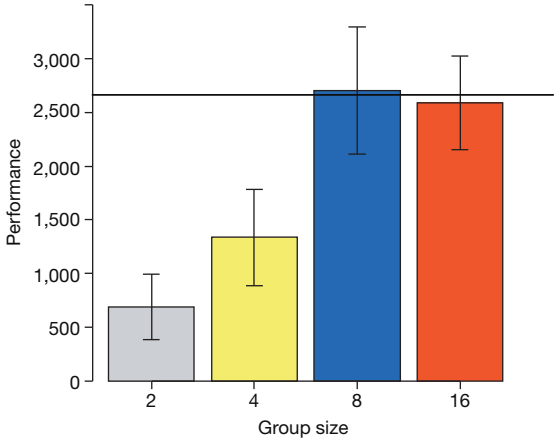


Figure 3 | Larger groups prevent degradation of the complex cultural trait. The horizontal line shows the fishing-net performance from the cultural package. Performance is measured using arbitrary life units. Plotted are the mean values ± s.e.m. The complex task deteriorated in the smaller groups (mean performance: 2-player groups = 685, $t = -6.50$, d.f. = 14, $P < 0.0001$; 4-player groups = 1,334, $t = -2.99$, d.f. = 11, $P = 0.01$) and remained stable in the larger groups (mean performance: 8-player groups = 2,706, $t = 0.07$, d.f. = 11, $P = 0.95$; 16-player groups = 2,590, $t = -0.17$, d.f. = 11, $P = 0.87$).

surprising, as the combination of inter-individual variance in cognitive abilities and sampling effect increase the probability of observing high performers within a large group. Furthermore, a group can collectively achieve a solution to a cognitive problem that is not available to an individual through 'swarm intelligence'[24]. Whatever the mechanism, the best within-group artefacts drove the performance of the entire group, as shown by the correlation between best within-group artefacts and other within-group artefacts at the final trial (arrowhead, Pearson correlation = 0.39, $t = 5.53$, d.f. = 167, $P < 0.0001$; fishing net, Pearson correlation = 0.29, $t = 2.78$, d.f. = 87, $P = 0.007$).

When technological complexity is measured by the number of existing tools in the cultural repertoire, archaeological data produce mixed results[6–9]. The occurrence of new tools is poorly understood, but individuals rarely invent new tools from scratch; pre-existing technologies should have a role through combination; that is, bringing together two established cultural traits to generate a new trait[18,25,26]. Interestingly, this game suggests that increasing group size favours the maintenance of cultural diversity, a prerequisite for subsequent innovation through combination. It is worthy of note that the aim of the game was to maximize the player's 'health'. Thus, a player not able to perform the complex task (for example, lacking good visual memory) could perform better by efficiently repeating the simple task than by trying the complex one. It suggests that the individual diversity associated with larger group size could be pivotal to the maintenance of cultural trait diversity. By facilitating the maintenance of cultural diversity, increasing group size could also favour the emergence of division of labour at the group level. Such conditions pave the way for the emergence of inter-individual collaborations and group-level organization, some of the most important properties of human groups[27].

At the individual level, results also show that complex-task (fishing net) copying was most of time associated with a loss of skill, whereas simple-task copying was not (see Supplementary Information). This confirms that greater loss of information is expected for a more complex task, as suggested by Henrich[4]. At the group level, the maintenance of the complex task observed in large groups is thus explained by an increased probability to observe rare events directly linked to group size, such as a perfect copy or even an innovation, rather than overall better individual copying abilities. Following an innovation, prestige bias leads individuals to shift, and copy a new model. Even if copying deteriorates information, the mean group performance can increase, allowing cultural evolution to operate[4]. Accordingly, cultural

complexity—as measured in the archaeological record, for example—is most probably not a direct marker of the mean cognitive ability, as an ecological increase in population size could trigger the onset of a cumulative cultural evolution. Such an event may subsequently lead to the evolution of advanced copying ability, as this trait will most probably be an advantage in such a cultural environment. The players' difficulty in properly copying the fishing net from the cultural package (100% of fishing-net builders failed at the first trial) also illustrates the importance of multiple demonstrations and multiple attempts in the acquisition process[28]. In our game, players acquired the correct skill over several trials. In large groups, high-performing copiers (more likely to be observed as group size increases) can prevent the skill from disappearing, enabling players who lack good copying ability to benefit from more demonstrations.

Our results support Henrich's hypothesis: changes in group size can generate both adaptive cultural evolution and maladaptive losses of culturally acquired skills[4]. In our evolutionary past, group-size reduction may have exposed human societies to notable risks, as humans live in many habitats to which they are ill-suited without specific cultural adaptations[11,12]. Indeed, the more that we depend for our survival on large bodies of culturally transmitted knowledge, the more we rely on living in large groups. Under such conditions, group-size reduction could have triggered important loss of skills, leading to societal collapse[13], particularly in challenging environments. Interestingly, some cumulative cultural innovations, such as writing, printing and various forms of long-term data storage, allow the preservation of information outside of individuals, such that it is unknown whether the maintenance of current cultural complexity is nowadays similarly dependent on group size.

METHODS SUMMARY

Each player was randomly assigned to a group of 2, 4, 8 or 16 players, and all groups started the game by benefiting from the same cultural package (composed of an arrowhead and a fishing net, see Methods section for the complete details of the game). The simple task involved drawing an arrowhead, for which the performance evaluation depended only on its shape. The arrowhead demonstration in the cultural package involved 15 steps and provided 1,638 life units. The complex task involved building a fishing net, for which the performance evaluation depended on its shape and the procedure used to build it. The fishing-net demonstration in the cultural package involved 39 steps (the sequence of which mattered) and provided 2,665 life units. The starting individual life level was 3,400 units, and 1,000 units (daily needs) were subtracted at each trial. The task difficulties were designed so that, for a non-experienced player, the probability of scoring below their daily needs (and thus having a negative score) was low when choosing the arrowhead task and high when choosing the fishing-net task. Each trial was followed by an information period during which players could choose a single demonstration to observe (ranked by their performance), from one of their group members or the cultural package. The cultural package was available up to the third trial: from the fourth trial and after, social information came only from players' group members. A total of 366 male students (mean age = 24.1 years, s.d. = 4.4) played this game only once, in groups of 2 (15 replicates), 4, 8 or 16 (12 replicates each) players.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 August; accepted 14 October 2013.
Published online 13 November 2013.

1. Boyd, R. & Richerson, P. J. *Culture and the Evolutionary Process* (Univ. of Chicago Press, 1985).
2. Tomasello, M., Kruger, A. C. & Ratner, H. H. Cultural learning. *Behav. Brain Sci.* **16,** 495–511 (1993).
3. Boyd, R. & Richerson, P. J. *Not by Genes Alone* (Univ. of Chicago Press, 2005).
4. Henrich, J. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *Am. Antiq.* **69,** 197–214 (2004).
5. Powell, A., Shennan, S. & Thomas, M. G. Late Pleistocene demography and the appearance of modern human behavior. *Science* **324,** 1298–1301 (2009).
6. Kline, M. A. & Boyd, R. Population size predicts technological complexity in Oceania. *Proc. R. Soc. B* **277,** 2559–2564 (2010).
7. Read, D. Population size does not predict artifact complexity: analysis of data from tasmania, arctic hunter-gatherers, and oceania fishing groups. *UC Los Angeles: Hum. Complex Systems* http://www.escholarship.org/uc/item/61n4303q (2012).
8. Collard, M., Kemery, M. & Banks, S. Causes of toolkit variation among hunter-gatherers: a test of four competing hypotheses. *Can. J. Archaeol.* **29,** 1–19 (2005).
9. Read, D. An interaction model for resource implement complexity based on risk and number of annual moves. *Am. Antiq.* **73,** 599–625 (2008).
10. Caldwell, C. A. & Millen, A. E. Human cumulative culture in the laboratory: effects of (micro) population size. *Learn. Behav.* **38,** 310–318 (2010).
11. Henrich, J. & Broesch, J. On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Phil. Trans. R. Soc. Lond. B* **366,** 1139–1148 (2011).
12. Boyd, R., Richerson, P. J. & Henrich, J. The cultural niche: why social learning is essential for human adaptation. *Proc. Natl Acad. Sci. USA* **108,** 10918–10925 (2011).
13. Diamond, J. The Tasmanians — the longest isolation, the simplest technology. *Nature* **273,** 185–186 (1978).
14. Boyd, R. & Richerson, P. J. Why does culture increase human adaptability. *Ethol. Sociobiol.* **16,** 125–143 (1995).
15. Boyd, R. & Richerson, P. J. Why culture is common, but cultural evolution is rare? *Proc. Br. Acad.* **88,** 77–93 (1996).
16. Tomasello, M. In *Social Learning in Animals: the Roots of Culture* (eds Heyes, C. M. & Galef, B. G.) 319–346 (Academic, 1996).
17. Tennie, C., Call, J. & Tomasello, M. Ratcheting up the ratchet: on the evolution of cumulative culture. *Phil. Trans. R. Soc. Lond. B* **364,** 2405–2415 (2009).
18. Lewis, H. M. & Laland, K. N. Transmission fidelity is the key to the build-up of cumulative culture. *Phil. Trans. R. Soc. Lond. B* **367,** 2171–2180 (2012).
19. Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B. & Laland, K. N. Identification of the social and cognitive processes underlying human cumulative culture. *Science* **335,** 1114–1118 (2012).
20. Derex, M., Godelle, B. & Raymond, M. Social learners require process information to outperform individual learners. *Evolution* **67,** 688–697 (2013).
21. Claidière, N. & Sperber, D. Imitation explains the propagation, not the stability of animal culture. *Proc. R. Soc. B* **277,** 651–659 (2010).
22. Diamond, J. *Guns, Germs, and Steel: the Fates of Human Societies* (W. W. Norton, 1999).
23. Marquet, P. A. *et al.* Emergence of social complexity among coastal hunter-gatherers in the Atacama Desert of northern Chile. *Proc. Natl Acad. Sci. USA* **109,** 14754–14760 (2012).
24. Krause, J., Ruxton, G. D. & Krause, S. Swarm intelligence in animals and humans. *Trends Ecol. Evol.* **25,** 28–34 (2010).
25. Basalla, G. *The Evolution of Technology* (Cambridge Univ. Press, 1988).
26. Henrich, J. & Boyd, R. On modeling cognition and culture: Why cultural evolution does not require replication of representations. *J. Cogn. Cult.* **2,** 87–112 (2002).
27. Smaldino, P. E. The cultural evolution of emergent group-level traits. *Behav. Brain Sci.* (in the press).
28. Flynn, E. & Whiten, A. Studying children's social learning experimentally ''in the wild''. *Learn. Behav.* **38,** 284–296 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Belkhir for help in establishing the Stuctured Qwery Language database, D. Dubois for recruiting participants and organizing the experimental sessions, and the Laboratory of Experimental Economics of Montpellier (University of Montpellier I) for hosting the experiment. Contribution ISEM 2013-146.

Author Contributions M.D., B.G. and M.R. designed the study. M.-P.B. and M.D. collected data. M.D., M.P.B. and M.R. analysed the data. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.D. (maxime.derex@gmail.com).

METHODS

Participants. A total of 366 male students were randomly selected from a database managed by the Laboratory of Experimental Economics of Montpellier (LEEM) and recruited by email from various universities in Montpellier (Southern France). The subjects ranged in age from 18 to 49 years (mean = 24.1 years, s.d. = 4.4 years). Each participant was randomly assigned to one condition of the experiment. Participants received fees for travel according to the LEEM operating rule (€2 for local participants, €6 for others).

Procedure. The experiment took place in a computer room at the LEEM. For each session, a maximum of 20 players sat at physically separated and networked computers and were randomly assigned to a group (the number of players per group varied according to the treatment, see below). They could not see each other, and they were blind with regard to the purpose of the experiment and who belonged to each group. The players were instructed that communication was not allowed. The participants could read instructions on their screens about the rewards and the goal of the game, and they were requested to enter their date of birth before the start of the game. At the end of the game, each subject received a reward according to his performance (€10 on average, see rewards calculation).

Principle. The participants played a computer game (programmed in Object Pascal with Delphi 6) during which they had to maximize their 'health' using two virtual tasks, making an arrowhead or a fishing net. Before the beginning of the game, players were advised that the fishing-net task was potentially more effective than the arrowhead task but that the fishing-net construction was more difficult. The participants were also informed that the performance of an arrowhead depended only on its shape, whereas the performance of the fishing net depended on its shape and the procedure used to build it. Each player began the game by observing a video demonstration of each task from a cultural package and was instructed that the arrowhead and fishing-net demonstrations could be improved. The arrowhead demonstration involved 15 steps and was associated with a score of 1,638. The fishing-net demonstration involved 39 steps (the sequence of which mattered) and was associated with a score of 2,665. The participants were not aware of the highest achievable score for any task.

The players then had 15 trials to collect resources and improve their health score. At each trial, they had the opportunity to build either an arrowhead or fishing net. Players began the game with a health score of 3,400 units. At each trial, their health level was reduced by 1,000 units, corresponding to their daily needs. Between trials, players could benefit from social information (see below).

Construction period. During the construction period (limited to 90 s), the players had to choose between the arrowhead task and the fishing-net task to collect resources.

The arrowhead task. The performance of an arrowhead depended only on its shape. The arrowhead score ranged from 0 to 2,400 units. A simple symmetric, triangular arrowhead constituted an acceptable performance equal to the player's daily needs. As a consequence, the probability of a non-experienced player scoring below his daily needs was low.

Construction details for the arrowhead task. First, the players had to choose the rectangular grid dimension on which to draw the arrowhead (30 possible values, Extended Data Fig. 1.a). Once the grid was chosen, the players had to draw their arrowhead. By clicking on the grid, the players could draw lines between points (Extended Data Fig. 1.b). The players had to draw the outline of their arrowhead and the virtual relief. No construction rules were implemented.

Score calculation for the arrowhead task. Once an arrowhead was drawn, it was evaluated by the program. The arrowhead was scanned pixel by pixel to evaluate five parameters: the size (α) and the symmetry (β) of the arrowhead, the number of notches (γ) and their regularities (δ), and the triangular shape (λ). All the parameters were compared to a theoretical optimal value and normalized from 0 to 1. The score S was then obtained according to this formula:

$$S = \alpha.400 + \beta.400 + \gamma.800 + \delta.400 + \lambda.400 \quad (1)$$

The fishing-net task. The participants had access to several virtual tools with which to build their nets. The performance of a net depended on its shape and the procedure used to build it. The net's score ranged from 0 to 5,135 units. Departure from the construction rules (which were unknown to the players) resulted in increased penalties during use of the fishing net. As a consequence, the probability of a non-experienced player scoring below his daily needs (1,000 units) was high.

Construction details for the fishing-net task. First, the players had to choose the squared grid dimension on which to build the net (30 possible values, Extended Data Fig. 1c). Once the grid was chosen, the players had access to different types of ropes and knots, as in a previous experiment[20]. A rope could be set between any pair of attaching points, and a knot could be tied to any attaching point, in any order (Extended Data Fig. 1d). There were limited ropes and knots available. Each additional rope placed on the frame decreased the length of the remaining rope according to the length used. This remaining quantity was visible on the screen. There were three different types of rope available (thick (red), medium (blue) and thin (green)). Each additional knot placed on the net decreased the length of the remaining knot quantity according to the type of knot used (three sizes available). This remaining quantity of knots was visible on the screen. Modification of one parameter produced complex interactions with others to generate a complex fitness landscape. For example, the use of the thickest ropes prevented the net from breaking but increased the net visibility so that the number of potentially caught fish was reduced. In addition, the order of construction (the process), was important. For example, two ropes that intersect at an attaching point should be tied together with a knot before another rope is put on the frame. If this step is omitted, the expected score is reduced.

Score calculation for the fishing-net task. Once a fishing net was constructed, it was evaluated by the program. A global resistance score (GR) was calculated according to the number of knots and compared to the required number. A local resistance score (LR_i) was determined for each mesh i according to the length and thickness of the ropes involved. During each virtual fishing exercise, 79 fish were launched, with a unique size of 65 (arbitrary units). The probability of each fish encountering the net increased according to the net overall size (set by the grid-spacing parameter) and decreased according to its visibility. The visibility of a net was computed as the sum of the lengths of all ropes used, weighted by their thicknesses. Once a fish was set to interact with the net, random coordinates were generated to identify at which mesh the interaction took place. If the fish was smaller than the mesh, it escaped. If it was larger, the probability of the net breaking was calculated as $1 - (GR*LR_i)$. In such a case, the whole fishing process stopped. If the net did not break, the fish could escape with a probability P_{esc}, which depended on the shape of the mesh and construction-rule penalty. If the fish did not escape, its size was added to the player's score. This process was repeated until the last fish was encountered or until the net broke.

Information period. After each trial, the resulting score, along with the player's health level, was displayed. The players could also see score lists for the arrowheads and fishing nets generated by the player's group members at the previous trial, ordered by performance. During the first three trials, the cultural package (arrowhead or fishing net) was included in the corresponding list.

By clicking on a score, the players could see the step-by-step procedure needed to build the selected item. Any demonstration lasted 40 s, regardless of the number of building steps. At each information period, a player could see only one demonstration. From the fourth information period, cultural-package demonstrations were removed from the lists. The players then had access only to their group member's demonstrations. The duration of the social-information period was 70 s.

Rewards calculation. The individual rewards were €10 on average. Players who died during the game (health level dropped below 0) earned €2. The other players earned an amount €A calculated according to this formula:

$$A = H_p/H_g.[5.N + 3.N_d] + 5 \quad (2)$$

where H_p is the player's health level, H_g is the sum of the group's health levels, N is the size of the group, and N_d is the number of dead players within the group.

Treatments. Four group sizes were considered: 2 players, 4 players, 8 players and 16 players. All treatments were replicated 12 times, except for the 2-player treatment, which was replicated 15 times.

Cultural evolution. The aim of the study was to investigate the evolution of the cultural packages that were introduced in the experimental groups. Two types of analyses were carried out; one examined the maintenance of cultural tasks (whether some individuals exploited the cultural task at the end of the game), and the other examined the performance associated with the tasks. For each of the two tasks, we focused on the best within-group information because this information drives subsequent cultural evolution (due to prestige bias).

Maintenance of cultural tasks. Two models were used. One model investigated how the simple task was maintained in comparison with the complex task. A cultural task was considered to be maintained within a group if, among the last three trials, at least one individual of the group exploited the task. The response variable was the presence or absence of each task in each group. The independent variables were the type of task (arrowhead or fishing net), group size, mean age within the group, and type of task × group size interaction, with 'group identity' as a random factor. Generalized linear mixed models (binomial) were used.

The other model investigated how cultural diversity was maintained according to group size. Cultural diversity was considered to be observed within a group if, in the last three trials, at least one individual performed the arrowhead task, whereas at least one individual performed the fishing-net task. The response variable was the presence or absence of the diversity. The independent variables were the group size and mean age within the group. A generalized linear model (binomial) was used.

Best within-group information. The performances of the best within-group arrowheads at the fifteenth trial were compared to the score of the arrowhead from the cultural package, using a one-sample Student's *t*-test (if the distribution significantly departed from normality, a Mann–Whitney-Wilcoxon test was also performed; results were qualitatively similar, data not shown). A further linear model was used to investigate the effect of group size. In this case, the response variable was the score of the best within-group arrowhead at the fifteenth trial, and the independent variables were group size and mean age within the group. These two analyses were carried out again for the fishing-net performances.

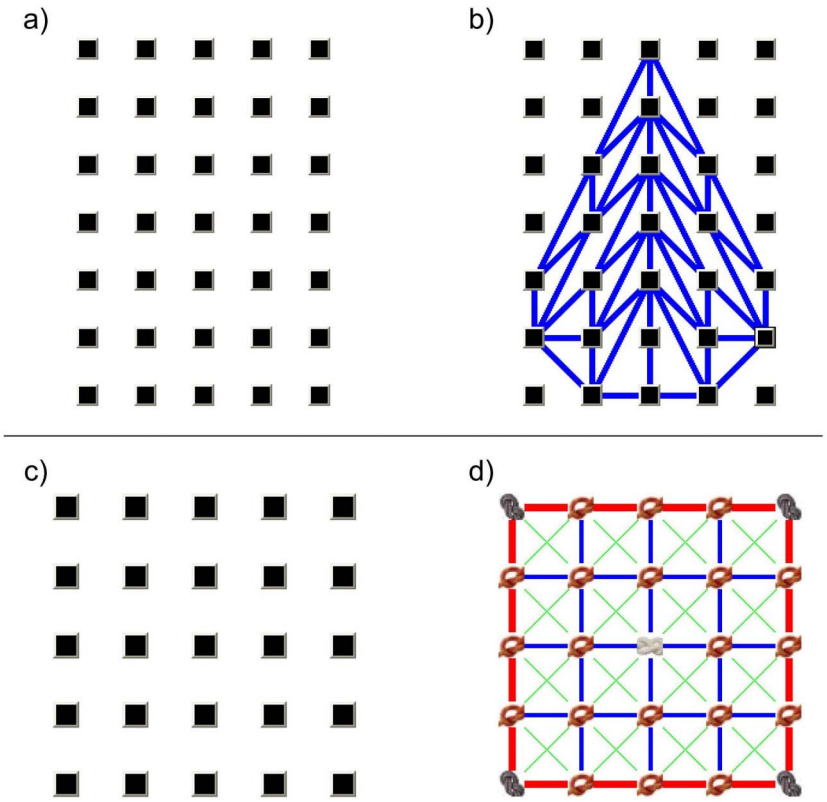
As groups could lose one of the two tasks, all analyses were carried out twice. In one case, we considered all groups, and performance score of zero was assigned when a task was lost from a group, that is, the degradation of the performance was considered complete (results shown in the main text). In the other case, we considered only the performance of the groups that conserved the task (results shown in Extended Data Figs 2 and 3).

Normality of residuals was significantly rejected (using Shapiro's test) in three models. This was owing to the presence of zero values (associated with task loss) generating a gap in the distribution between zero and the minimal score. When the presence or absence of the task was explicitly controlled for in order to estimate this gap, normality of residuals were not rejected (sometimes requiring the exclusion of only one outlier). All results described here were unchanged, whether or not these changes were made.

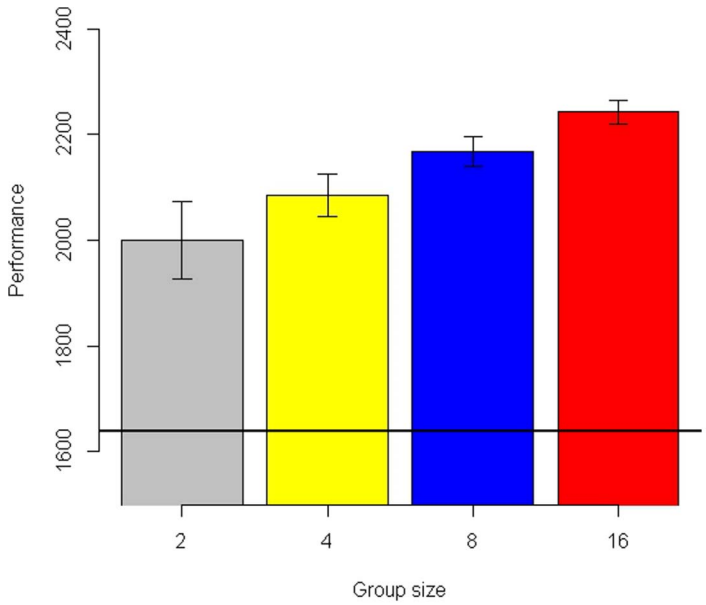
Fidelity of copying. Henrich's model assumes that information transmission is generally imperfect (particularly with complex tasks). Indeed, if copying is faithful, no cultural losses are expected. For each task, analyses were carried out to evaluate copying fidelity. During the observation period, players could choose a single demonstration to observe before building a new artefact. The aim was to study whether or not artefacts built by the players performed worse than the artefacts they observed.

An observed artefact was considered as a model and was associated with n copies, depending on how many players observed the same model. For example, if three players observed the same model, three copies (copy 1, copy 2 and copy 3) were created. All possible pairs of artefacts were formed from the model and the copies: with one model and three copies, this corresponded to 6 pairs (model–copy 1; model–copy 2; model–copy 3; copy 1–copy 2; copy 1–copy 3; copy 2–copy 3). Comparisons of 'model–copy' represent our treatment of interest: if copying deteriorates information, the expected score difference (model score minus copy score) should be positive (null or negative otherwise). Comparisons of 'copy–copy' represent a control treatment: the expected score difference should be null. The focal artefact (first artefact from the pair) was either a copy or the model and was always compared to a copy (second artefact from the pair). The skill was considered to have deteriorated when the focal artefact outperformed the copy (score difference strictly positive). The binary response variable was the presence or absence of skill degradation. The independent variables were the type of the focal artefact ('copy' or 'model'). The identity of the focal artefact and the identity of the producer of the second artefact from the pair were included as random effects. A generalized linear mixed model (binomial) was used. All analyses were carried out separately for each task (arrowhead and fishing net).

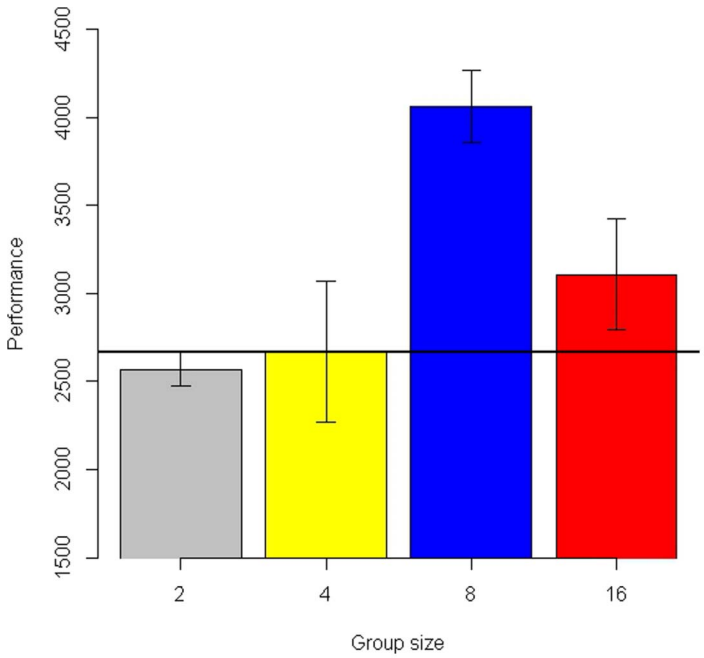
Correlation between best within-group information and individual performances. This study was culture-centred, focusing on the state of the information available within groups (how the best within-group information performed). Considering that the best-within-group information influences the subsequent performance of the entire group, it is important to test the correlation between best within-group information and individual performances: owing to prestige bias, the best within-group information should affect the performance of the entire group. We examined the correlation between the best within-group information and the performance of the other players at the fifteenth trial using the Pearson correlation test.



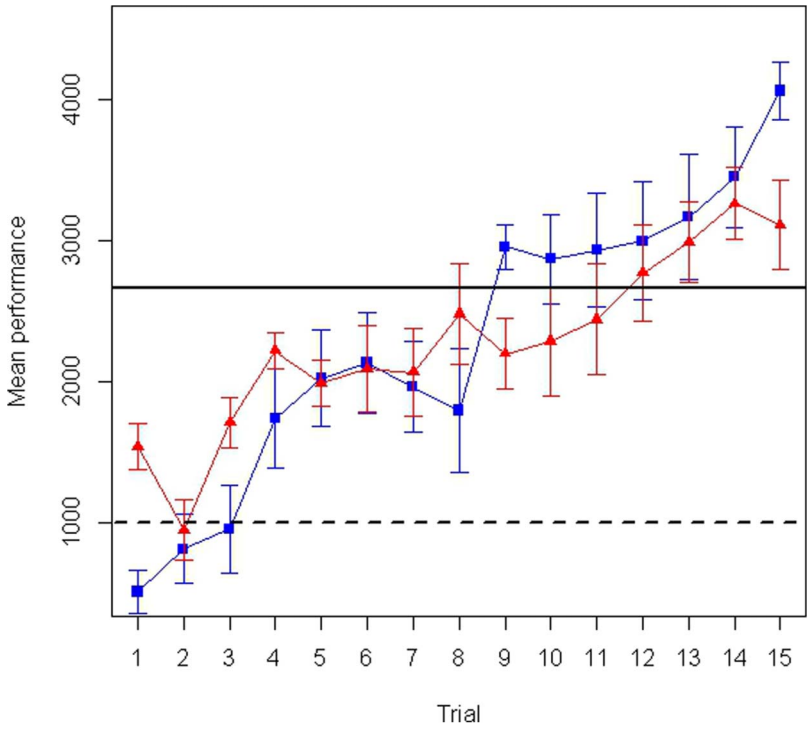
Extended Data Figure 1 | Cultural tasks. **a**, Rectangular grid composed of 35 attaching points in which to draw an arrowhead. The spacing between the attaching points was modifiable. **b**, An example of an arrowhead. **c**, Square grid composed of 25 attaching points in which to build a fishing net. The spacing between the attaching points was modifiable. **d**, An example of a fishing net.



Extended Data Figure 2 | Best within-group information associated with the simple task, when conserved within the group. Performance is measured using arbitrary life units. Plotted are the mean values ± s.e.m. Considering only the performance of the groups that conserved the task (see Methods), the simple task of the cultural package was improved in all group sizes (mean performance: 2-player groups = 2,000, $t = 4.90$, d.f. = 10, $P = 0.0006$; 4-player groups = 2,085, $t = 11.12$, d.f. = 8, $P < 0.0001$; 8-player groups = 2,166, $t = 18.84$, d.f. = 11, $P < 0.0001$; 16-player groups = 2,242, $t = 27.57$, d.f. = 11, $P < 0.0001$). Group size had a linear effect on the performance of the best within-group arrowhead ($F_{1,41} = 15.3$, $P = 0.0003$). The horizontal line shows the performance of the arrowhead from the cultural package.



Extended Data Figure 3 | Best within-group information associated with the complex task, when conserved within the group. Performance is measured using arbitrary life units. Plotted are the mean values ± s.e.m. Only 4 2-player groups (26.7%) conserved the complex task and were therefore excluded from the analysis. The complex task was stable in the 4-player groups (mean performance = 2,669, $t = 0.01$, d.f. = 5, $P = 0.99$) and improved in the larger groups. The difference between 8-player groups and the demonstration of the cultural package was significant (mean = 4,059, $t = 6.79$, d.f. = 7, $P = 0.0001$, one-sided) but marginally significant concerning 16-player groups (mean = 3,108, $t = 1.40$, d.f. = 9, $P = 0.09$, one-sided). Group size had a linear and an unexpected quadratic effect on the performance of the best within-group fishing net ($F_{1,24} = 10.6$, $P = 0.003$ and $F_{1,24} = 9.88$, $P = 0.004$, respectively). This quadratic effect could indicate that participants had trouble making use of the information in a large group, but our experimental design allows us to rule out this possibility (see Supplementary Information). Instead, early performances of 16-player groups affected the probability of observing the cultural-package demonstration, hindering players from acquiring pivotal information (see Extended Data Fig. 4 and Supplementary Information). The horizontal line shows the performance of the fishing net from the cultural package.



Extended Data Figure 4 | Best within-group information associated with a fishing net (when conserved within the group) across time. The red line shows 16-player groups and the blue line shows 8-player groups. Performance is measured using arbitrary life units. Plotted are the mean values ± s.e.m. At the beginning of the game, the 16-player groups performed better than the 8-player groups ($F_{1,22} = 21.7$, $P = 0.0001$), as expected. However, the opposite was observed at the end of the game ($F_{1,16} = 5.68$, $P = 0.03$). During the first three trials, the performance associated with the best within-group fishing net affected the probability of observing the cultural-package demonstration. Thus, the probability of observing the cultural-package demonstration was lower in 16-player groups compared with 8-player groups. A lower rate of observation of the cultural-package reduced the group performance suggesting that the observation of demonstrations from other sources hindered the acquisition of pivotal information (see Supplementary Information for details). It suggests that, under specific conditions, the increasing number of valuable sources of information associated with larger group size could lead to a suboptimal cultural evolution rate. The horizontal solid line shows the performance of the fishing net from the cultural package. The horizontal dashed line shows the players' daily needs.