nature biotechnology

Digital-to-biological converter for on-demand production of biologics

Kent S Boles[1,2], Krishna Kannan[1,2], John Gill[1], Martina Felderman[1], Heather Gouvis[1], Bolyn Hubby[1], Kurt I Kamrud[1], J Craig Venter[1] & Daniel G Gibson[1]

Manufacturing processes for biological molecules in the research laboratory have failed to keep pace with the rapid advances in automization and parellelization[1–3]. We report the development of a digital-to-biological converter for fully automated, versatile and demand-based production of functional biologics starting from DNA sequence information. Specifically, DNA templates, RNA molecules, proteins and viral particles were produced in an automated fashion from digitally transmitted DNA sequences without human intervention.

The manufacturing industry has continuously evolved new business models. A principal example is distributed manufacturing, which decentralizes the manufacturing process such that a robust supply chain coordinates production of a final product from intermediates manufactured at geographically distant locations[1–3]. Distributed manufacturing also refers to on-demand, small-scale production by consumers (close to the intended destination), such as the RepRap initiative for 3D printing of 3D printers[4,5]. Rapid progress in synthetic biology and biotechnology methods to synthesize biologics has, however, outpaced development of improved manufacturing technologies.

Approaches that reduce production timelines and simplify manufacturing processes for biological material are needed[6]. For example, demand-based synthesis of biological materials, including higher-order assemblies such as synthetic chromosomes[7,8] and viral particles[9], could accelerate the design–build–test cycle for research. In biomedical applications, emergency responses to disease outbreaks in areas lacking robust infrastructures could be achieved by need-based manufacturing of therapeutics at the site of use[10–12]. Reliable, portable manufacturing units could also facilitate the delivery of personalized medicine and vaccines[10–13].

We present a scalable bio-manufacturing unit, called a digital-to-biological converter (DBC), that receives digitally transmitted DNA sequence information and converts it into biopolymers, such as DNA, RNA and proteins, as well as complex entities such as viral particles, without any human intervention (**Fig. 1**, **Supplementary Figs. 1** and **2** and **Supplementary Videos 1**–**4**). The DBC integrates multiple software and instruments through a system controller (Hudson Robotics). This includes Archetype, a proprietary software (Synthetic Genomics, Inc.) that designs an optimized assembly strategy for input DNA sequence and splits it into overlapping oligonucleotides (**Supplementary Code**); SoftLinks V (Hudson Robotics) for the underlying workcell automation; an oligonucleotide synthesizer operated by Poseidon software (BioAutomation) that initiates the synthesis of oligonucleotides automatically; robotic parts for oligonucleotide post-synthesis processing and accurate liquid handling, orchestrated by SoloSoft software (Hudson Robotics); and an automated thermocycler that functions as a DNA assembly module and production platform for biopolymers synthesis.

Although no new code or software was written specifically for the DBC, several hardware parts were configured for each run to accommodate differences in the DNA amplicons synthesized and the methods for activating the amplicons through transcription, translation or transformation to produce the final biological material. Comprehensive information about the software and the underlying configuration files used in each run as well as equipment and their interconnectivity is detailed in Online Methods (**Supplementary Figs. 1** and **2**).

We set out to demonstrate production of a series of biological materials without human intervention on the DBC. First, we synthesized a 1.5-kb DNA fragment encoding GFP. We manually designed a DNA template (GenBank KY199424) and entered the sequence into Archetype, which automatically converted input sequence into oligonucleotide sequences in a file written to a specified folder on the local network. The DBC was loaded with consumables, and the configuration files (Online Methods and **Supplementary Code**) specific for producing the DNA encoding GFP were designated in the DBC system controller (Online Methods). Upon detection of the oligonucleotide file (**Supplementary Table 1**), the DBC system controller transferred the file to the MerMade 192R oligonucleotide synthesizer (Online Methods, component 1 in **Supplementary Figs. 1**,**2** and **Supplementary Video 1**). Once synthesis was complete, the oligonucleotide-containing columns were moved by the lower crane arm (**Supplementary Fig. 1**, component 5) to a vacuum nest. An ammonium hydroxide–methylamine (AMA) deprotection solution was added to remove protecting groups from the oligonucleotides into a deep-well block and elute the oligonucleotides from the columns (Online Methods and **Supplementary Fig. 1**, components 10 and 11). For this, the oligonucleotide-containing block was sequentially

[1]Synthetic Genomics, Inc., La Jolla, California, USA. [2]These authors contributed equally to this work. Correspondence should be addressed to D.G.G. (dgibson@syntheticgenomics.com).

Received 5 August 2016; accepted 24 March 2017; published online 29 May 2017; doi:10.1038/nbt.3859

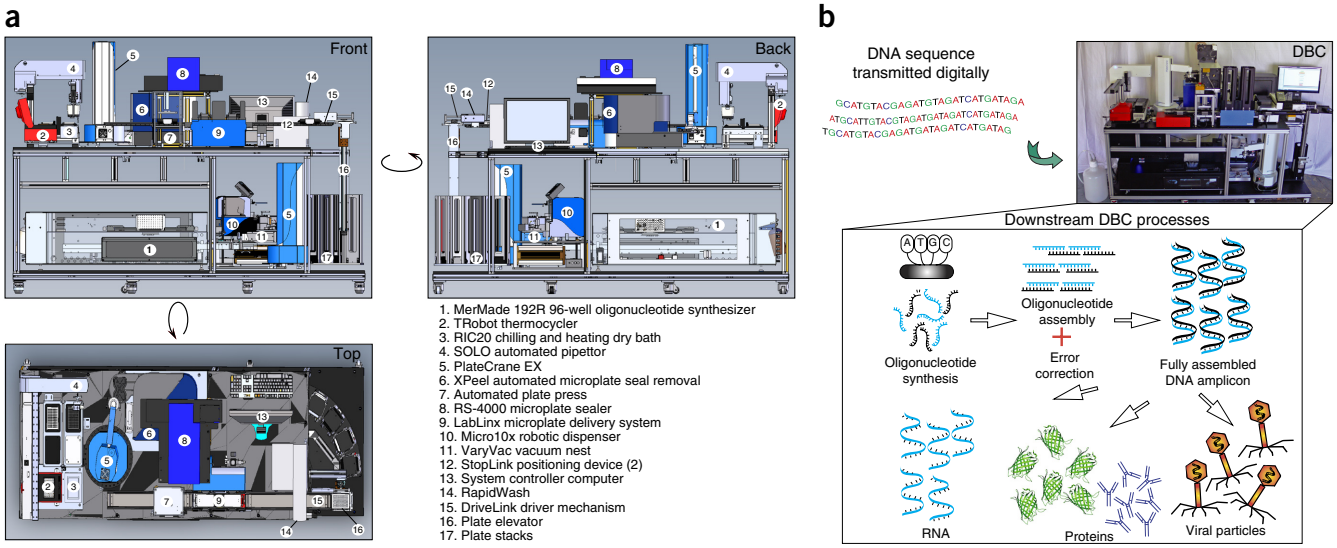

Figure 1 Concept of on-demand production of biologicals on the DBC. (**a**,**b**) Features (**a**) and applications (**b**) of the DBC. DNA sequence information received automatically initiates conversion of DNA sequence into constituent oligonucleotide sequences and feeds this information to an oligonucleotide synthesizer. Synthesized oligonucleotides are pooled, assembled, error corrected and reassembled into DNA amplicons within a thermocycler. DNA amplicons then direct the production of RNA, proteins and viral particles.

transferred by the lower plate crane, the elevator LabLinx, and the upper plate crane through components 5, 16 and 5 (**Supplementary Fig. 1**) to an automated plate sealer (**Supplementary Fig. 1**, component 8 and **Supplementary Video 2**). After sealing, the block was transferred to a press (**Supplementary Fig. 1**, component 7), using the upper crane arm and LabLinx (**Supplementary Fig. 1**, components 5 and 9), for 2 h to complete the AMA incubation. In a similar fashion, the block was moved and unsealed on an automated plate peeler (**Supplementary Fig. 1**, component 6) and returned by the LabLinx to the plate washer (**Supplementary Fig. 1**, component 14), which was modified to blow air out of the dispense needles into the block so that the AMA solution could be evaporated after deprotection of the oligonucleotides. The block was subsequently transferred by LabLinx and plate crane to the deck of an automated pipetting system (**Supplementary Fig. 1**, component 4) that could access the block in addition to an automated thermocycler and a dry bath held at 4 °C (**Supplementary Fig. 1**, components 2 and 3) preloaded with reagents necessary for gene assembly, error correction and *in vitro* transcription and translation (IVTT) (Online Methods).

The oligonucleotides were pooled from the individual wells of the block to a single well by the automated pipetting system (**Supplementary Video 3**). This pool was pipetted into a PCR plate on the thermocycler, where it was serially diluted through the entire first column. A PCR mix for gene assembly was pipetted from the dry bath into each of the oligonucleotide pool dilutions, and a PCR cycle was performed to assemble the target amplicon. Subsequently, samples were pipetted from the first column to the next, and error correction mix was added and incubated. Error correction ensured that a relatively high fraction of accurate DNA molecules were synthesized on the DBC (**Supplementary Fig. 3** and **Supplementary Table 2**). Next, the PCR mix for post-error-correction recovery was pipetted into each well of the next column, to which the error-corrected DNA was added, and a PCR cycle was initiated (**Supplementary Video 4**). Last, after the recovery PCR, the IVTT mix was pipetted into the adjoining column, and the assembled amplicon was added and incubated to synthesize GFP. The optimization of protein production conditions on the DBC is described in Online Methods (**Supplementary Fig. 4**). This constituted the end of the automated synthesis run on the DBC.

Following the DBC run, functionality of the DBC-synthesized GFP was evaluated by measuring fluorescence (**Supplementary Fig. 5**). The GFP-encoding amplicon synthesized on the DBC was visualized by agarose gel electrophoresis, and the protein was visualized by SDS–PAGE (**Fig. 2a**). The consensus sequence of the final PCR amplicon was verified using Sanger sequencing (**Supplementary Data**). In some instances, the amplicons were cloned into plasmids and transformed into *Escherichia coli*, and DNA synthesis error rates were assessed by sequencing 48 or 96 individual clones (**Supplementary Table 2**). Specifically, for the amplicons encoding GFP, hemagglutinin (HA) and neuraminidase (NA), >70% of the molecules, on average, were found to be error-free at the end of the recovery PCR following error correction (**Supplementary Table 2**).

We next synthesized Orencia (abatacept), Lucentis (ranibizumab) and Herceptin (trastuzumab) antibody polypeptides[14–16] (GenBank KY199428, KY199429 and KY199430, respectively) on the DBC using a similar method (**Fig. 2b**, **Supplementary Fig. 6**, **Supplementary Tables 3–5** and Online Methods). The polypeptides were detected on SDS–PAGE using the FluoroTect Green$_{Lys}$ labeling system (Promega), and the amplicons were sequence verified by Sanger sequencing (**Supplementary Data**). These antibodies are not expected to be functional because the *E. coli* lysates that we used for transcription and translation do not promote proper folding of antibodies.

We next applied the DBC to produce an RNA vaccine and a bacteriophage, both of which have potential as therapeutics for infectious diseases. We chose to synthesize a Venezuelan equine encephalitis virus (VEEV) sequence, because VEEV has been reported to function as an effective RNA-based vaccine vector[17,18]. We synthesized an empty VEEV backbone lacking the single open reading frame (ORF) encoding the five VEEV structural genes, into which a vaccine antigen gene of choice could be readily inserted. Prior to the DBC run, we manually inserted an HA coding sequence from influenza A (H7N9 bird flu) to produce an HA-antigen-expressing RNA replicon as a

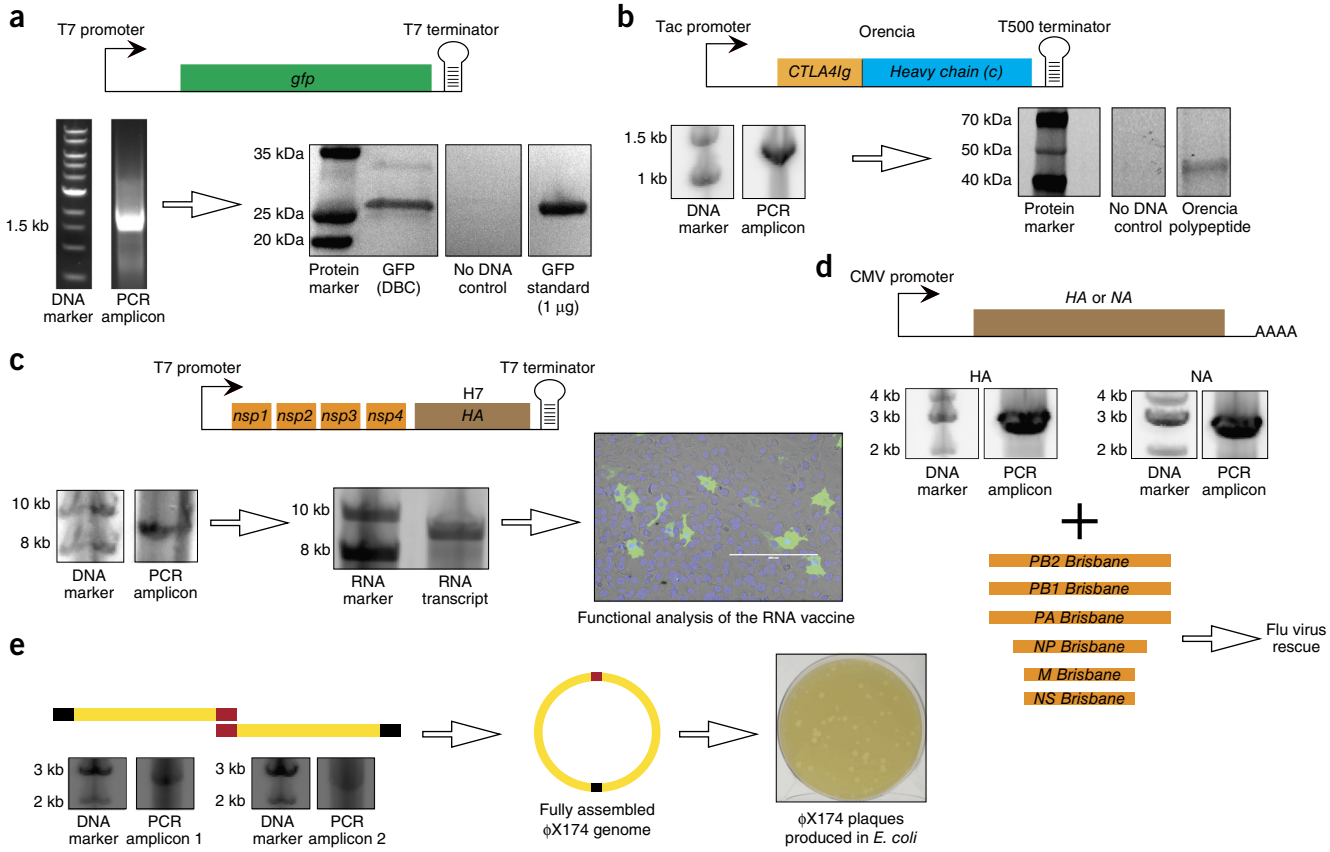

Figure 2 Biologicals production on the DBC. (**a**) GFP production. Top, DNA construct; bottom, visualization of amplicon by gel electrophoresis (left) and of protein product by Flamingo fluorescent gel stain (right). (**b**) Antibody-constituting polypeptide synthesis. Top, Orencia construct with translationally fused light (*CTLA4Ig*) and heavy chains; bottom, visualization of amplicon by gel electrophoresis (left) and detection of polypeptides using FluoroTect Green$_{Lys}$ (right). (**c**) RNA vaccine production and functional analysis. Top, VEEV-based replicon backbone[18]. Bottom, visualization of DNA amplicon (left), RNA transcript of the replicon (middle) and overlay of H7, DAPI and brightfield channels (right). (**d**) Flu virus rescue. Top, DNA sequences encoding HA and NA with control elements; bottom, visualization of HA and NA amplicons and schematic of co-transfection with plasmids encoding backbone viral genes (Brisbane, B/Brisbane/60/2008,Victoria lineage) into MDCK cells to rescue viral particles (polymerase basic protein 2 (PB2), polymerase basic protein 1 (PB1), polymerase acidic protein (PA), nucleoprotein (NP), matrix 1 (M1), and nonstructural protein 1 (NS1)). (**e**) Phage production. Left, genome of the *E. coli* phage ΦX174 synthesized as two overlapping PCR fragments, assembled into a circular genome (middle), transformed into the NEB-10β *E. coli* cells on DBC and assayed offline for functionality by further infecting *E. coli* str. HF4704 (right). Parts of the DNA, RNA and protein gels are displayed; full-length gels are shown in **Supplementary Figure 10**.

positive control (GenBank KY199425, **Supplementary Table 6** and Online Methods). We then adapted the methods for GFP production on the DBC to synthesize an H7 HA amplicon and used Gibson assembly on the DBC to clone it into the empty VEEV backbone[17]. We also transcribed this H7-carrying replicon on the DBC using an *in vitro* transcription reaction mix (Online Methods). Following automated synthesis on the DBC, the transcription reaction carrying the RNA vaccine was recovered and shown to be functional by transfection into Vero cells and immunostaining with an anti-H7 HA antibody (**Fig. 2c** and **Supplementary Figs. 7** and **8**).

We also produced functional influenza viral particles (H1N1) using genes encoding HA and NA synthesized on the DBC[18,19] (**Fig. 2d**). Briefly, HA and NA amplicons were synthesized on the DBC using methods similar to those outlined for GFP production. In addition, a pool of presynthesized oligonucleotides (Integrated DNA Technologies) encoding an upstream promoter and a downstream terminator sequence was included during the gene assembly. At the end of the DBC run, HA and NA amplicons were made, carrying the necessary control regions for expression in mammalian cells (GenBank KY199426 and KY199427; **Supplementary Tables 7** and **8**). These amplicons, together with plasmids for the remaining six influenza genes, were transfected into Madin–Darbyn canine kidney epithelial (MDCK) cells according to the protocol described by Dormitzer *et al.*[19] (Online Methods). Sanger sequencing verified the HA and NA amplicons and the control regions, and viral titers of up to 9.01×10^7 PFU/ml at day 6 demonstrated that functional influenza viral particles were produced (**Supplementary Tables 2** and **9**).

Finally, we fully automated production of the ΦX174 bacteriophage, which has a 5,386-bp genome, on the DBC. The genome sequence was manually designed *in silico* as two overlapping pieces (30-bp overlap; GenBank KY199431 and KY199432; **Supplementary Table 10**) and each piece was synthesized on the DBC[13]. These two amplicons were joined via Gibson assembly on the DBC to produce a fully circularized genome that was transformed into chemically competent *E. coli* cells using the thermocycler present in the DBC (**Supplementary Table 11** and Online Methods). After the DBC run, we confirmed by plaque assay that the ΦX174 particles were functional (~100 plaques obtained from 200 µl lysate) and verified the sequence identity of the synthesized genome using Sanger sequencing (**Fig. 2e** and **Supplementary Data**).

The DBC is a prototype that demonstrates the technical feasibility of completely automated, on-demand manufacturing technology for biological research and production. Some features of the current DBC design, such as the DNA assembly processes, have been incorporated into BioXp 3200 (SGI-DNA), a commercially available, automated DNA assembly system.

The DBC in its present form can produce a high percentage (>70%) of error-free DNA molecules (calculated by individual cloning of three different amplicons) (**Supplementary Table 2**). We synthesized DNA amplicons of different lengths, GC contents and generated a variety of end products from these DNA templates (**Supplementary Fig. 9** and **Supplementary Table 12**). The MerMade 192R synthesizer can generate up to 192 oligonucleotides during a single DBC run, which can be used to synthesize a 6-kb DNA fragment or multiple smaller fragments. This capacity is large enough to accommodate most bacterial genes or ORFs encoding human therapeutic proteins such as antibodies; thus, the DBC in its present form is primed for on-site generation and testing of DNA constructs (and the RNA or protein products thereof) without the extended times required for manufacturing when external vendors are involved.

One limitation of the current DBC is its dimensions and consequently, portability. To reduce the size and number of robotic parts for liquid handling and oligonucleotide synthesis, microfluidic and microarray technologies could be used to substitute large portions of several DBC components. Further, this would enable an improved scale of DNA synthesis, which would in turn enable additional DBC functionality such as high-throughput functional screening of proteins (enzymes and antibodies) by combinatorial assembly of expression constructs using randomized parts. In addition, the error rates of the amplicons produced on the DBC could be improved by the integration of novel error-correction technologies (improved enzymes and methodologies).

Although the antibody polypeptides we produced were nonfunctional, the incorporation of cell-free lysates that assist correct folding and post-translational modification have the potential to enable production of functional antibodies and other human therapeutic proteins from sequence information[20]. Incorporation of additional analytical instruments such as spectrophotometers or electrophoretic separation systems would allow automated measurement of yields and monitoring of product quality.

In conclusion, the DBC prototype fully integrates and automates processes from oligonucleotide design and synthesis to the production of biopolymers. Development of a smaller and portable DBC could enable reliable production at the point of demand and potentially reduce costs and increase access to bio-production in research laboratories. Finally, with the incorporation of large-scale synthesis technologies, one can envision the DBC being used in industrial settings to enable high-volume production of biologics such as proteins and RNA vaccines.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank C. Beard, S. Farah, O. Fetzer, K. Han, C. Hutchison, A. Lee, D. Lomelin, A. Nandi, T. Newman-Lehman, T. Peterson, S. Riedmuller, J. Robinson, H. Smith, J. Strauss, J. Thielmier, L. Warden, M. Winstead, and A. Witschi for their contributions to this work, and the US Defense Advanced Research Projects Agency (contract HR0011-13-C-0073 for D.G.G. and J.C.V.) for funding aspects of this work.

AUTHOR CONTRIBUTIONS

J.C.V. and D.G.G. conceived the study; K.S.B., K.K., J.G., H.G., B.H., K.I.K. and D.G.G. designed experiments and analyzed data; K.S.B., K.K., M.F. and D.G.G. performed experiments; and K.S.B., K.K., J.C.V. and D.G.G. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Rauch, S., Dallinger, M., Dallasega, P. & Matt, D.T. Sustainability in manufacturing through distributed manufacturing systems (DMS). *Procedia CIRP* **29**, 544–549 (2015).
2. Srai, J.S. *et al.* Distributed manufacturing: scope, challenges and opportunity. *Int. J. Prod. Res.* **54**, 6917–6935 (2016).
3. Pearce, J.M. Materials science. Building research equipment with free, open-source hardware. *Science* **337**, 1303–1304 (2012).
4. Jones, R. *et al.* RepRap—the replicating rapid prototyper. *Robotica* **29**, 177–191 (2011).
5. Pearce, J.M., Anzalone, N.C. & Heldt, C.L. Open-source wax RepRap 3-D printer for rapid prototyping paper-based microfluidics. *J. Lab. Autom.* **21**, 510–516 (2016).
6. Gottschalk, U., Brorson, K. & Shukla, A.A. The need for innovation in biomanufacturing. *Nat. Biotechnol.* **30**, 489–492 (2012).
7. Gibson, D.G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).
8. Hutchison, C.A. III *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).
9. Smith, H.O., Hutchison, C.A. III, Pfannkoch, C. & Venter, J.C. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. USA* **100**, 15440–15445 (2003).
10. Perez-Pinera, P. *et al.* Synthetic biology and microbioreactor platforms for programmable production of biologics at the point-of-care. *Nat. Commun.* **7**, 12211 (2016).
11. Pardee, K. *et al.* Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248–259.e12 (2016).
12. Li, J. *et al.* Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221–1226 (2015).
13. Khan, K.H. DNA vaccines: roles against diseases. *Germs* **3**, 26–35 (2013).
14. Moreland, L., Bate, G. & Kirkpatrick, P. Abatacept. *Nat. Rev. Drug Discov.* **5**, 185–186 (2006).
15. Narayanan, R., Kuppermann, B.D., Jones, C. & Kirkpatrick, P. Ranibizumab. *Nat. Rev. Drug Discov.* **5**, 815–816 (2006).
16. Hudis, C.A. Trastuzumab—mechanism of action and use in clinical practice. *N. Engl. J. Med.* **357**, 39–51 (2007).
17. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
18. Geall, A.J. *et al.* Nonviral delivery of self-amplifying RNA vaccines. *Proc. Natl. Acad. Sci. USA* **109**, 14604–14609 (2012).
19. Dormitzer, P.R. *et al.* Synthetic generation of influenza vaccine viruses for rapid response to pandemics. *Sci. Transl. Med.* **5**, 185ra68 (2013).
20. Stech, M.K.S. & Kubick, S. Cell-free synthesis meets antibody production: a review. *Antibodies* **4**, 12–33 (2015).

ONLINE METHODS

General operational procedure on the DBC. The DBC was designed as two decks to decrease the footprint (**Fig. 1** and **Supplementary Fig. 1**). The MerMade 192R synthesizer is connected by an RJ45 cable to a Realtek PCIe GBE Family Controller Gigabit Ethernet network interface card installed in the system controller computer on the top deck. The remaining instruments on the DBC are controlled through serial cables (RS-232) routed to a Digi Edgeport 416 Serial Adaptor (USB to 20 serial ports) attached to the system controller computer. The lower deck holds a BioAutomation Mermaid 192R synthesizer capable of presenting and accepting two 96-well DNA synthesis chucks simultaneously. It is loaded and unloaded by a PlateCrane EX (Hudson Robotics) with four plate stacks. The Micro10X Robotic Dispenser (Hudson Robotics), proximal to the PlateCrane, is attached to a VaryVac Vacuum Nest (Hudson Robotics) such that a 96-well block from a plate stack can be loaded to the Vacuum Nest and a 96-well DNA synthesis chuck can be placed on top of it so that when the Micro10X dispenses elution/deprotection solution into the chuck, the vacuum nest can draw a vacuum and pull the solution through the synthesis columns into the synthesis chuck. Once the chuck has been returned to the plate stack for storage, the block can be moved to the top deck of the DBC by a Plate Elevator (Hudson Robotics). Once at the top level, the Plate Elevator has a DriveLink drive mechanism that pushes the block onto the track of the LabLinx Microplate Delivery System. A RapidWash microplate washer is installed such that it can service a plate on the LabLinx track. Furthermore, the Lablinx has a press integrated over the track to put pressure on the blocks during deprotection incubation. The LabLinx itself carries two plate stacks. One is used to store the blocks containing oligonucleotides as needed, and the other is used to dispense pipet tips boxes. On the opposite end of the LabLinx from the elevator is a second PlateCrane EX (Hudson Robotics) that allows the block to be loaded on to a RS-4000 Microplate Sealer (Brandel) that is physically located on top of a XPeel Automated Microplate Seal Removal (Brooks) system, where the PlateCrane can also place the block. Furthermore, a SOLO Automated Pipettor (Hudson Robotics) is within range of the PlateCrane so that it can move the blocks and pipet tip boxes from the LabLinx onto its deck. Next to the SOLO are a TRobot Thermocycler (Biometra GmbH) and a Ric20 Chilling and Heating Dry Bath (Torrey Pines Scientific, 15 × 1.5-ml tube block) oriented and located in such a way that the SOLO can directly pipet solutions between these instruments. A schematic of different parts of the DBC and the sequence of operations is shown in **Supplementary Figure 2**.

Oligonucleotide synthesis on the DBC. The design of the entire DNA sequence containing all the necessary features for assembly and expression under appropriate conditions was done manually. Codon optimization was done wherever necessary using the program JCat[21]. This DNA sequence was pasted into our proprietary Archetype automated oligonucleotide design program, which generated a CSV file carrying the oligonucleotides necessary for the synthesis of the product that is written to a specific network folder. The script 'Oligonucleotide Designer' (executable by WordPerfect), that was used for generating oligonucleotides from DNA sequence is available at https://github.com/kentboles/DBC and also in the **Supplementary Code**.

The DBC system controller (Hudson Robotics) software is comprised of an interface written in Visual Basic that allows the user to define run characteristics such as the name of the run, oligonucleotide synthesis script and the cycling conditions for gene assembly. Once the run has been initiated, the software monitors a specific network folder for the arrival of the necessary CVS file. Once present, the software parses the oligonucleotides into a simple text format that can be read by the Poseidon software operating the MerMade 192R synthesizer (BioAutomation) in a 2 × 96-well format (an automated version of the 192E). The oligonucleotide file, synthesis script and other user-defined inputs are transferred to a mySQL database until called. The system controller launches SoftLinks V (Hudson Robotics) that runs sub-routines specific for oligonucleotide synthesis, oligonucleotide deprotection, pooling and gene assembly. The oligonucleotide synthesis routine loads two 96-well DNA synthesis chucks preloaded with synthesis columns into the instrument and transfers the oligonucleotide file and synthesis script to the MerMade synthesizer, launching a run.

After completion of oligonucleotide synthesis, the first synthesis chuck was robotically -transferred to the Micro10x and VaryVac by a PlateCrane EX (Hudson Robotics) where the oligonucleotides were eluted into a 96-well, deep-well (2-ml) block with 300 µl of equal parts ammonium hydroxide and methylamine (AMA solution). Specifically, 100 µl was added three times, each with an initial pulse of vacuum (15 inches Hg for 5 s) followed by a 5-min incubation and a final vacuum to pull the AMA solution into the deep-well block (25 inches Hg for 5 min). The block was then lifted by a plate elevator (Hudson Robotics) to the top deck of the DBC and transferred by the LabLinx Microplate Delivery System (Hudson Robotics) and a PlateCrane EX to a RS-3000 plate sealer (Brandel) where it was sealed with foil tape. It was then incubated under a press integrated into the LabLinx for 2.5 h. Subsequently, the block was transferred by the LabLinx and PlateCrane EX to an XPeel Automated Microplate Seal Removal instrument (Brooks Automation). The unsealed block was then automatically transferred back to the LabLinx Microplate Delivery System's integral plate stack. The second block was moved to the VaryVac, where the deprotection process was similarly conducted up until the incubation under the press. While the second block was being pressed, the first block was transferred to a RapidWash (Hudson Robotics), modified to inject air (100 psi) into each well of the block for 4 h to speed up the evaporation of the base (from deprotection). Subsequently, the block was transferred to LabLinx Microplate Delivery System's integral plate stack. Following this, the second block in the press, was similarly peeled and injected with air on the RapidWash for 4 h. Both blocks were then transferred to the pipetting platform using the PlateCrane (Hudson Robotics). 10 µl of TE (pH 8.0) was used to suspend each oligonucleotide. These were further combined to obtain a pool of oligonucleotides. The pool was transferred using the SOLO Automated Pipettor (Hudson Robotics) to well A01 of a PCR plate in the TRobot thermocycler (Biometra GmbH). 10 µl 1:1 dilutions of the pool were made in wells B01-H01 to reach a target concentration (typically, 25 nM) that is ideal for gene assembly. Of note, the design of components 2–5 of the DBC (**Supplementary Fig. 1**) were used in the development of the BioXp 3200 prototype. The run times on the DBC were ~16 h to generate a full plate (96 wells) of oligonucleotides, ~11 h for deprotection and drying of the plates, ~11 h for oligonucleotide pooling, gene synthesis and error correction and recovery PCR reactions, plus the time required for additional downstream processes (e.g., transcription, coupled transcription and translation, restriction digestion, Gibson assembly[17], transformation). Oligonucleotide synthesis by amidite chemistry is an imperfect process. Correct base coupling is highly reliant on incubation times and the absence of moisture during the synthesis process. The MerMade 192R synthesizer (BioAutomation) was operated with various synthesis scripts before a final protocol was selected. Starting with a synthesis script provided by BioAutomation (DBC-1.scp), we reduced the amounts of most of the reagents and increased the coupling time to arrive at a synthesis script that was sufficiently optimized to produce predominantly error-free DNA molecules ((DBC_Final.scp). Comprehensive configuration files and scripts for SoftLinks V, the synthesizer, SOLO pipettor, and TRobot thermocycler are available at https://github.com/kentboles/DBC and also in the **Supplementary Code**. Operations on the DBC are also demonstrated in **Supplementary Videos 1–4**.

DNA assembly on the DBC. The procedure described below was carried out by the DBC (unless mentioned otherwise) in a stepwise and automated fashion. DNA assembly was done using a method similar to that described by Hutchison *et al.*[8], with modifications to adapt the method to the DBC. Briefly, 10 µl oligonucleotide mix (synthesized and pooled on the DBC) was added to 45 µl of Q5 High-Fidelity DNA Polymerase (NEB) carrying 0.5 µM of each of the distal primers and 0.04% PEG-8000 (the mix was premade and loaded on to the DBC). A PCR assembly cycle of 98 °C for 1 min, 30× (98 °C for 30 s, 65 °C for 3 min with further extension of 15 s per cycle) and 72 °C for 5 min was performed. Error correction was performed by denaturing and slow annealing to allow the formation of heteroduplexes of error-carrying DNA molecule with the 'correct' DNA molecule. Specifically, denaturation at 98 °C for 2 min and slow annealing to 85 °C and at a rate of 2 °C/s and incubation at this temperature for 2 min; slow annealing was further carried out by cooling the mix to 25 °C at a rate of 0.1 °C/s and incubating for 2 min at 25 °C. Heteroduplexes were then cleaved and eliminated by addition of 2.7 µl PCR reaction to 5.3 µl water, 2 µl Surveyor Nuclease (IDT) and 1 µl exonuclease III (NEB) (diluted 1:4,000) and incubated at 42 °C for 1 h. A recovery PCR was

performed to amplify the error-corrected DNA. Conditions for the recovery PCR were identical to the initial assembly reaction, with the exception that 2.5 µl of error-corrected DNA was added to 47.5 µl of the master mix with Q5 High-Fidelity DNA Polymerase (NEB). Following the final PCR, all DNA templates synthesized on the DBC (GFP, polypeptides constituting antibodies (Orencia, Lucentis, Herceptin), HA H1N1, NA H1N1, HA H7 (H7N9), ΦX174 genome) were recovered manually and verified to be 100% accurate by consensus Sanger Sequencing (**Supplementary Data**).

GFP, HA and NA amplicons synthesized on the DBC (before and after error correction) were manually cloned into pUC19 vector using Gibson assembly and transformed into TransforMax EPI300 electrocompetent cells (Epicentre). Plasmids were manually extracted from 48 or 96 colonies, and the inserts were sequenced using Sanger sequencing. Error rates were determined in two ($n = 2$) or three replicates ($n = 3$), and the results, including percentage correct clones, number of insertions, deletions, base mutations and total error rate, are reported as averages and s.d. and s.e.m. in **Supplementary Table 2**. As a comparison, GFP amplicon was synthesized manually with oligonucleotides from IDT using the DNA assembly process employed on the DBC. This amplicon (before and after error correction) was also cloned and sequenced, and error rates were determined using the above process.

Optimization of protein and RNA synthesis methods for the DBC. In its current setup, the DBC does not have a provision for the incubation of reagents at −80 °C, which is recommended by the manufacturers as the optimal temperature to maintain reagents until the start of the transcription–translation reaction. Hence, the following experiments were done offline from the DBC to overcome this potential issue and identify optimal conditions for use on the DBC. We tested the PURExpress *in vitro* Protein Synthesis Kit (New England BioLabs) and the *E. coli* S30 Extract System for Linear Templates (Promega) kit by preparing the premix of the respective transcription–translation reactions without the DNA template and incubating this mix for 16 h at −20 °C or 4 °C before the addition of the DNA template. This mimics the 'waiting' time on the DBC while the DNA template is synthesized and assembled. DNA template was subsequently added to this mix, and the transcription–translation reaction was carried out (**Supplementary Fig. 4**). We also attempted to drive the *in vitro* transcription–translation reaction using PCR product directly after assembly on the DBC without further purification. We initially tested equal amounts of the column-purified and unpurified T7-GFP template in the PURExpress system. Surprisingly, the unpurified template resulted in <15% loss of activity compared to the purified template (**Supplementary Fig. 4**). We observed a similar effect when we tried the T7- RiboMAX Large Scale RNA Production System (Promega) for RNA production and the *E. coli* S30 Extract System for Linear Templates (Promega) for protein production. We further explored the tolerance limit of the unpurified PCR to drive translation in PURExpress *in vitro* Protein Synthesis Kit by adding different volumes of the unpurified template into the reaction (**Supplementary Fig. 4**). The use of unpurified DNA template to drive subsequent reactions obviates the need for additional reagents or instrumentation for DNA purification on the DBC. We also identified that the inclusion of T7-terminator improves the yield of GFP; hence, the terminator was included in our constructs wherever necessary (**Supplementary Fig. 4**).

Cell-free transcription and translation on the DBC. To produce GFP on the DBC, the oligonucleotides described in **Supplementary Table 1** were synthesized and subjected to the assembly process described above. Following the final recovery PCR, 5 µl unpurified PCR amplicon was directly added to 20 µl premade PURExpress mix (NEB) automatically by the DBC. The transcription–translation mix was manually prepared according to manufacturer's instructions (with an additional 20U of SUPERase In RNase Inhibitor (ThermoFisher Scientific)) and preloaded on the DBC before the start of the run. This mix was held at 4 °C on the DBC throughout oligonucleotide synthesis and DNA template assembly (>24 h). The translation reactions were manually recovered after a 3-h incubation at 37 °C and manually processed offline to visualize the expressed protein and test functionality. Specifically, the reactions were filtered using a 100-kDa cutoff filter (Amicon, EMD Millipore) to remove the ribosomes in the PURExpress mix, and the filtrate was treated with 1–10 µg RNase A (ThermoFisher Scientific) for 15 min at 37 °C. Following this, the entire reaction was precipitated with four volumes of 100% ice-cold acetone and resuspended directly in 50 µl 1X loading buffer (Bio-Rad). Samples were resolved using a 4–15% Mini-PROTEAN TGX Precast Protein Gel (Bio-Rad). To quantify the GFP produced, we also loaded different amounts of control GFP standard protein (Vector Laboratories) alongside the sample obtained from the DBC. Gels were stained using the Flamingo Fluorescent Gel Stain (Bio-Rad) and visualized on a Typhoon FLA 9500 Imager according to the manufacturer's instructions (**Fig. 2a** and **Supplementary Fig. 10a**). GFP synthesized on the DBC was verified to be functional by measuring fluorescence (480 nm excitation and 510 nm emission) (**Supplementary Fig. 5**).

For the production of polypeptides constituting the antibodies, we were unsuccessful in producing detectable amounts of the polypeptides in the PURExpress system (NEB) during the initial offline tests. Hence, we used the FluoroTect GreenLys *in vitro* Translation Labeling System (Promega) in the *E. coli* S30 Extract System for Linear Templates (Promega). 3.5 µl unpurified PCR amplicon (carrying *tac* promoter) was directly added to 21.5 µl *E. coli* S30 Extract System for Linear Templates prepared according to manufacturer's instructions and preloaded on to the DBC before the start of the run. This transcription–translation mix carried 20 U of SUPERase In RNase Inhibitor (ThermoFisher Scientific) and 1 µl FluoroTect GreenLys tRNA (Promega). Following a 3-h incubation at 37 °C, which is the final step of the DBC run, the reaction was manually recovered and processed offline using the following method. Reaction was treated with 1–10 µg RNase A (ThermoFisher Scientific) for 15 min at 37 °C, precipitated with four volumes of 100% ice-cold acetone and resuspended directly in 50 µl 1X loading buffer (Bio-Rad). Samples were resolved using a 4–15% Mini-PROTEAN TGX Precast Protein Gel (Bio-Rad) and visualized using Typhoon FLA 9500 Imager according to the manufacturer's instructions (**Fig. 2b** and **Supplementary Figs. 6** and **10b**). Because the antibody-polypeptides were synthesized using *E. coli* cell lysates, they do not undergo post-translational modifications that are essential for folding and function of the antibodies. Hence, the functionality of these polypeptides was not determined; only the synthesis was monitored using the above-described method.

Production of RNA-based replicon vaccine encoding H7 antigen on the DBC. A potential vaccine candidate, the hemagglutinin coding sequence from an influenza A (H7N9 bird flu, Shanghai_2_2013) was synthesized from oligonucleotides made by the DBC after transmission of the design files from Archetype encoding the 1,790-bp ORF (**Supplementary Table 6**). Subsequent to the oligonucleotide synthesis and assembly, the H7 coding region was further assembled into an RNA replicon template cassette (9,601 bp)[19] on the DBC through Gibson assembly. The self-amplifying RNA replicon is based on the genome of the Venezuelan equine encephalitis virus (positive-strand RNA), which, similarly to a eukaryotic mRNA, carries a 5′ cap and a polyadenylation motif. Upon access to the cytosol of a mammalian cell, the nonstructural genes (NSP1–4) are transcribed and the RNA genome is replicated by making a negative strand resulting in a double-stranded RNA genome. This double-stranded genome encodes a 26S subgenomic promoter that drives the production of structural proteins for viral replication. We used the viral genome without the structural genes to produce the H7 antigen. Following the H7 assembly on the DBC, 1 µl PCR was added to Gibson assembly mix along with equimolar amounts of the arms of RNA replicon backbone in a 10-µl reaction on the DBC. The self-replicating RNA molecule backbone was amplified with the following primers: 5′For, 5′-GCGGCCGCCTTTCGTTTTATTTGACCATGTTGGTATG-3′; 5′Rev, 5′- GGTGCGATATCTTGGCGGACTAGACTATGTCGTAGTC-3′; and 3′For, 5′- TTAATTGATCGATACAGCAGCAATTGGCAAGCTGCTTACATAG-3′; 3′Rev, 5′-CCTGCAGGCCTAACTGTGGCCAGTCCAGTTACGCTG-3′ to produce the segments 5′ and 3′ to the H7 ORF. The assembly reaction was incubated at 50 °C for 3 h on the DBC. 1 µl assembly reaction was transferred to a 100-µl recovery PCR with Q5 High-Fidelity DNA Polymerase (NEB) along with 0.5 µM each of the distal primers For, 5′-GACCATGTTGGTATGTAATACGACTCACTATAGATAGGCGGCGCATGAGAGAAG-3′ and Rev, 5′-CTGCAGGCCTAACTGTGGCCAGTCCAGTTACGCTGAAAAAAACC-3′ on the DBC. The PCR was cycled at 98 °C for 1 min, 30× (98 °C for 30 s, 72 °C for 5 min 30 s), 72 °C for 10 min as final extension step. 5 µl PCR reaction was directly transferred by the DBC to the T7-RiboMAX Large Scale RNA Production System (Promega) reaction

doi:10.1038/nbt.3859

mix. This reaction mix included 0.9 mM final concentration of rGTP instead of the regular 7.5 mM concentration to allow for the capping of the mRNA, 20 U of the SUPERase In RNase Inhibitor (ThermoFisher Scientific) and the Ribo m^7G Cap Analog (Promega) at 3 mM final concentration. The transcription reaction was incubated for 4 h at 37 °C. Following the transcription reaction on the DBC, the transcripts were manually recovered and purified offline. Reaction was stopped by the addition of RQ1 DNase (Promega) and incubation at 37 °C for 15 min. Transcripts were purified using the RNeasy mini kit (Qiagen) and 1 µg was transfected into Vero cells using Lipofectamine 2000 following the manufacturer's suggested method. After transfection, cells were incubated in an atmosphere of 5% CO_2 at 37 °C for 12 h. Cells were fixed and permeabilized with methanol–acetone (1:1) for 5 min at room temperature. Cells were then immunostained with anti-H7 primary antibody (ImmuneTech; 1:250) at 37 °C for 1 h, then washed with 1× PBS and incubated with Alexa Fluor 488–conjugated secondary antibody (Life Technologies; 1:1,000) at 37 °C for 1 h. For DNA staining, cells were incubated with DAPI. Images were acquired using an EVOS digital inverted-fluorescence microscope. The staining was compared to cells transfected with a bench-made replicon expressing H7 and mock controls. A part of the RNA sample used for transfection was reverse transcribed to DNA with a primer downstream of the H7 gene (**Supplementary Fig. 7**) and Sanger sequenced (**Supplementary Data**) to demonstrate the H7 portion of the amplicon was 100% as designed.

Influenza virus rescue using DNA parts synthesized on the DBC. DNA encoding hemagglutinin (HA) or neuraminidase (NA) from isolate A/Brisbane/10/2010 EPI278594 (H1N1) was synthesized on the DBC from oligonucleotides synthesized and assembled on the machine (**Supplementary Table 7**). Each gene was flanked by a CMV promoter at the 5′ end and a bovine growth hormone–derived polyadenylation signal at the 3′ end. These regulatory elements were pooled manually from presynthesized oligonucleotides (IDT) and added to the pool of HA and NA nucleotides synthesized on the DBC at 25 nM concentration during the assembly (**Supplementary Table 8**). The procedure for DNA assembly of the HA and the NA genes along with the regulatory elements was performed on the DBC as described above. The PCR reactions were subsequently transferred to Novartis (Holly Springs, North Carolina), where the H1N1 virus was successfully rescued offline from the DBC according to the protocol described by Dormitzer *et al.*[19]. Briefly, 1 µg of each linear HA or NA synthetic cassette was co-transfected into MDCK 33016PF cells along with 1 µg of each ambisense plasmid that encodes PA, PB1, PB2, NP, NS, or M influenza genes. The cells were incubated at 37 °C and 4% CO_2 for 6 and 8 d, and the supernatant was assayed for viral particles. Viral particle titers in the cell culture medium were calculated by transferring tissue culture supernatant to MDCK cell monolayers followed by immunostaining for influenza NP (Millipore)[19].

Production of phage ΦX174 on the DBC. The 5.3-kb genome of the phage ΦX174 (GenBank NC_001422.1) was divided *in silico* using Archetype into two segments overlapping at both ends (60-bp overlap at TCCTTTGAAT GGTCGCCATGATGGTGGTTATTATACCGTCAAGGACTGTGTGACT ATTGA) to enable *in vitro* homology–based circularization of the genome (**Supplementary Table 10**). Each segment was flanked by a 30-bp sequence followed by a NotI restriction site to enable amplification using the following primers: forward, 5′-CCTTTCGTTTTATTTGACCATGTTGGTATGGCGGC CGC-3′; reverse, 5′-GGCCTAACTGTGGCCAGTCCAGTTACGCTGGCGGC CGC-3′. The 30-bp additional sequence was not related to the ΦX174 genome, but was included to enable amplification of the two genomic segments: 5′ end including the NotI site, CCTTTCGTTTTATTTGACCATGTTGGTATGGC GGCCGC; 3′ end including the NotI site, GGCCTAACTGTGGCCAGTCCA GTTACGCTGGCGGCCGC. Two segments of the ΦX174 were synthesized from oligonucleotides generated by Archetype, synthesized and assembled on the DBC. General methodology of assembly was as described above, but the PCR mix carried the corresponding terminal primers. Following assembly, the two DNA segments were digested with 20 U NotI for 1 h at 37 °C on the DBC to expose the homologous ends of the two subgenomic fragments. After digestion, the two fragments were ligated using Gibson assembly at 50 °C for 1 h on the DBC to obtain circular genomes in a 10-µl reaction. The entire assembly reaction was transformed into 100 µl NEB-10β cells by heat shock at 42 °C for 4 s on the TRobot thermocycler (on the DBC), after which the transformed cells were cooled to 4 °C and maintained at that temperature for 30 min on the DBC. Details of the optimization of this transformation procedure on the DBC are listed in **Supplementary Table 11**. This mixture was recovered by the addition of 2 ml KCl medium[9] on the DBC. Following this, the transformed cells were manually recovered from the DBC and incubated at 37 °C for 45 min. Transformed cells were then manually lysed, and ΦX174 viral particles were tested for functionality by assaying their ability to produce plaques when plated with a susceptible *E. coli* C-strain, HF4704, as described by Smith *et al.*[9]. Approximately 100 plaques were observed when 200 µl lysate was plated from one of the genome assemblies, which is lower than the previously observed efficiency (~180 plaques per 100 µl volume of the lysate)[9]. Control plates were negative for plaques, and the DNA amplicons from the two half-genomes were Sanger sequenced to show they matched the target sequence 100% (**Supplementary Data**).

Statistics. Error rates were determined by direct counts of Sanger sequencing of cloned amplicons encoding GFP, HA or NA from two or three separate synthesis cycles on the DBC. Influenza rescue experiments were performed in duplicate. Data are averages, s.d. and s.e.m.

Data availability. Data are available from the corresponding authors upon reasonable request.

21. Grote, A. *et al.* JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* **33**, W526–W5331 (2005).