**Partitioning of a 2-bit hash function across 65 communicating cells**

Jai Padmakumar1 and Christopher A. Voigt1

1 Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

Potential reviewers (round 1):

Round 2:

\*Correspondence and request for materials should be addressed to C.A.V. (cavoigt@gmail.com)

**Key words:** Synthetic Biology, Systems Biology, genetic circuit design automation, cryptography, bitcoin, blockchain, computer aided design, Boolean logic

Supplementary figures to make

To do

* Circuit datasheets
* Gate datasheets
* ~~ECF response functions~~
* ECF induction curve
* ~~ECF genetic schematic~~
* ~~Marionette induction data~~
  + ~~Data made~~
* ~~Full response functions for pLux, pCin, pPhlF, and pRpa~~
  + ~~Data made~~
* ~~Sender:receiver response functions with data points~~
  + ~~Data made~~
* ~~Cross-talk bar plots with data points~~
* ~~Cross-talk schematic~~
* Cell gating for stationary phase and exponential phase
* ~~Rpa precursor optimization~~
  + ~~Data made~~
* ~~Example calculation of 1 iteration of md5~~
* ~~Full MD5 before partitioning minimization~~
  + ~~Figure made~~
* ~~Minimized md5 after partitioning minimization~~
  + ~~Figure made~~
* ~~Readthrough with N15 example~~
  + ~~Data in illustrator (note mention axis are plasmid rpu)~~
* Other repressors that didn’t end up in final set?
* Strains table list
* DNA sequence tables
* Yosys synthesis file



**Supplementary Figure 1: Demonstration of gate readthrough** Response functions of gates with the repressor and output in different orientations are shown. Gates are based on the N15 repressor and are carried on p15a plasmids. The gate is insulated using a strong double terminator, DT42. Despite that, a noticeable increase in measured promoter activity is observed (middle plot) due to readthrough of the downstream DT42 terminator at high induction levels of the PTet-PTac tandem promoter. When the repressor and output positions are swapped (left plot) or when the output is put in the opposite direction, the uptick in promoter activity is not observed. Cells were grown overnight in M9 media containing 50ug/uL kanamycin, diluted 1:100 and cultured for 2 more hours, before being diluted 1:625 for 4 hours with appropriate inducers. After 4 hours, cells were analyzed by flow cytometry and data was converted to plasmid RPUs using the same method to convert to genome RPU but with the reference standard carried on a p15a plasmid. Points represent independent replicates taken on different days (single replicate for middle plot, 2 replicates for left and right plots).

****

**Supplementary Figure 2: Schematic of cross-talk constructs** A total of 144 constructs assaying cross-talk between repressors and promoters were cloned. Repressors were cloned into the attB7 landing pad while all promoters were cloned into the attB2 landing pad. Cross-talk was measured by measuring output fluorescence after 0uM and 1000uM IPTG induction.

****

**Supplementary Figure 3: Raw cross-talk data** The bar plots show the raw data for the heatmap. Repressorswere cloned into the attB7 landing pad and promoters were cloned into the attB2 landing pad of JAI\_MKC300. Cells were then induced with 0uM (gray bars) or 1000uM (blue bars) IPTG and the output was converted to RNAP/s (materials and methods). Data points represent independent replicates taken on different days.



**Supplementary Figure 4: Validation of marionette sensor array in JAI\_MKC300** Expression from nine of the marionette sensors in the landing pad strain (JAI\_MKC300) is shown. Induction values were 0uM (gray bars) and the maximum inducer concentration for each sensor (blue bars). Maximum inducer values were as follows: 25uM DAPG (JAI\_MKC334), 100uM Cuma (JAI\_MKC335), 10uM Lux (JAI\_MKC322), 100uM Van (JAI\_MKC336), 1000uM IPTG (JAI\_MKC337), 200nM aTc (JAI\_MKC338), 4000uM Ara (JAI\_MKC340), 100uM Sal (JAI\_MKC342) and 10uM OHC12 (JAI\_MKC323). Reporter promoters were described in Meyer et al 2018 except an alternative PTet promoter was used. Promoters were as follows: PPhlF,PCymRC, PLuxB, PVanCC, PTac, PTet, PBAD, PSalTTC, and PCin. As with the NOT gates, outputs were measured from the attB2 landing pad. Cells were grown overnight in M9 media, diluted 1:100 for 1.5 hours, then diluted 1:2000 with appropriate inducers and cultured for 5 hours. After 5 hours, fluorescence was analyzed by flow cytometry and converted to RNAP/s (Materials and Methods). Points represent independent replicates taken on different days.



**Supplementary Figure 5: Response functions for receiver cells alone** Response functions for the four receiver cells used in the sender:receiver devices are shown. As before, cassettes were integrated into the attB2 landing pad of JAI\_MKC300. Note that the *luxR*, *cinRAM*, and *phlFAM* sensors are contained within the Marionette sensor array while *rpaRAM*is not. Induction conditions for Lux receiver cells (top left) were 0uM, 0.0001uM, 0.0005uM, 0.001uM, 0.005uM, 0.01uM, 0.05uM, 0.1uM, 0.5uM, 1uM, 5uM, and 10uM OC6. For Cin receiver cells (top right), induction conditions were 0uM, 0.0001uM, 0.001uM, 0.005uM, 0.01uM, 0.05uM, 0.1uM, 0.25uM, 0.5uM, 1uM, 5uM, and 10uM OHC12. For Rpa receiver cells (bottom left), induction conditions were 0nM, 0.005nM, 0.01nM, 0.1nM, 0.25nM, 0.5nM, 1nM, 5nM, 10nM, 25nM, 50nM, and 100nM pC-HSL with 100nM of precursor p-Coumarate added to the media. For DAPG receiver cells (bottom right), induction conditions were 0, 0.0001uM, 0.001uM, 0.01uM, 0.1uM, 0.25uM, 0.5uM, 1uM, 2uM, 5uM, 10uM, and 20uM DAPG. Cells were cultured overnight in M9 media, diluted 1:1000 into appropriate inducer, and cultured for 3 hours. Fluorescence was measured by flow cytometry and converted to RNAP/s (Materials and Methods).



**Supplementary Figure 6: Response functions for sender:receiver devices** Complete response functions for the sender:receiver devices in Fig 2D are shown. Plots show response functions for OC6 (top left), OHC12 (top right), pC-HSL (bottom left), and DAPG (bottom right). Sender cells were characterized using a PTac promoter with the following IPTG values for induction: 0uM, 10uM, 20uM, 30uM, 40uM, 50uM, 70uM, 100uM, 150uM, 200uM, and 1000uM. As before, cassettes were integrated into the attB2 landing pad of JAI\_MKC300. Characterization was performed as previously described (Materials and Methods). Points represent independent replicates taken on different days.



**Supplementary Figure 7: Optimization of p-Coumarate concentration for pC-HSL production** Production of pC-HSL with varying concentrations of the precursor p-Coumarate is shown. Three NOR gates (columns) using repressor JR12 with different input promoters were tested with 0nM (bottom row), 10nM (middle row), and 100nM (top row) of the precursor were assayed with the previously described circuit characterization protocol (Materials and Methods). Ultimately, we chose to use 100nM of the precursor in subsequent experiments.



**Supplementary Figure 8: Detailed example of 2-bit MD5 calculations** The MD5 algorithm requires 4 rounds of 16 iterations each to hash a 32-bit input message, outputting an 8-bit hash. Our implementation performs 1 such iteration. The initial message is divided into 16 2-bit chunks, each of which is fed into the algorithm in a particular sequence for a given round[cite original algorithm]. There are 8 2-bit inputs to each iteration. Mj represents one 2-bit chunk of the message. Ai, Bi, Ci, and Di are variables that contain the final hash values and are updated at each step of the iteration (i indicates iteration number). T and S are constants that are pre-determined by the user. At each iteration, 6 different operations are performed. The first operation in the process (giving Y1) is a non-linear equation that changes each round: F() in round 1, G() in round 2, H() in round 3, and I() in round 4. All other calculations are equivalent. The schematic on the right visually depicts the operations and input/outputs (red box denotes modulo), as well as how the variables A, B, C, D are updated. After 64 iterations, the final values of A, B, C, and D are concatenated to produce the final 8-bit hash. Note the initial values for A, B, C, D, S, and T are supplied by the user and must be held constant to produce the same hash for the same input. In the original 32-bit MD5, the values were empirically determined for producing optimal hashing. In addition, the initial input message is padded to make the message 32-bit longs (padded bits denoted as “b”). This step, and the final concatenation step, are not performed by our circuit.



**Supplementary Figure 9: Partially minimized MD5 circuit diagram with gate numbers** The MD5 circuit diagram after logic synthesis, initial minimization, and partitioning is shown (Materials and Methods). Shaded boxes represent circuits assigned to a single cell. Gates not in shaded box represent cells containing only that single gate. As in Figure 4A, colored lines represent different sender signalling molecules. The circuit contains a total of 65 cells and 131 gates.



**Supplementary Figure 10: Fully minimized MD5 circuit diagram with gate numbers** The fully minimized MD5 circuit after subcircuit minimization and mapping to OR gates, as shown in Figure 4a, is shown here with the same gate numbers as the unminimized circuit. As before, shaded boxes represent circuits in individual cells, gates not in a shaded box represent cells containing only a single gate, and colored lines represent different sender signalling molecules. The circuit shown here contains 110 gates and 65 cells.



**Supplementary Figure XX: Schematic of ECF20 based gates** A) ECF20 is an orthogonal sigma factor that binds the E. coli sigma70 RNA polymerase core. Once bound, it directs transcription from ECF20-dependent promoters, which contain unique -35 and -10 sequences. B) Schematic of an example construct using Gate JR1-j1R3:P\_JR1-134 (JAI\_MKC204). Sigma70-dependent promoters are shown in green while the output ECF20-dependent promoter is shown in orange. The design is similar to that used for sigma70 based NOT gates, except ECF20 is cloned into the attB7 landing pad with an inducible promoter to induce ECF20 expression.



**Supplementary Figure XX: Response functions of ECF20 gates** Response functions of ECF20 based gates are shown. The gates were profiled using the same protocol for the sigma70 based gates (materials and methods) except 10uM OHC14 was added to induce expression of ECF20. Subplot titles show the strain number, repressor, RBS, and ECF20-based promoter. Constructs were tested in *E. coli* YJP\_MKC173. Points represent independent replicates taken on different days.



**Supplementary Figure XX: Representative flow cytometry** gates A) Representative gating of stationary phase *E. coli*. B) Representative gating of exponential phase *E. coli*.