Method

OPEN ACCESS

molecular systems biology

In situ genotyping of a pooled strain library after characterizing complex phenotypes

Michael J Lawson†, Daniel Camsund†, Jimmy Larsson, Özden Baltekin, David Fange & Johan Elf*

Abstract

In this work, we present a proof-of-principle experiment that extends advanced live cell microscopy to the scale of pool-generated strain libraries. We achieve this by identifying the genotypes for individual cells *in situ* after a detailed characterization of the phenotype. The principle is demonstrated by single-molecule fluorescence time-lapse imaging of *Escherichia coli* strains harboring barcoded plasmids that express a sgRNA which suppresses different genes in the *E. coli* genome through dCas9 interference. In general, the method solves the problem of characterizing complex dynamic phenotypes for diverse genetic libraries of cell strains. For example, it allows screens of how changes in regulatory or coding sequences impact the temporal expression, location, or function of a gene product, or how the altered expression of a set of genes impacts the intracellular dynamics of a labeled reporter.

Keywords DuMPLING; live cell; microfluidic; single cell; strain libraries
Subject Categories Methods & Resources; Quantitative Biology & Dynamical Systems
DOI 10.15252/msb.20177951 | Received 23 August 2017 | Revised 11 September 2017 | Accepted 22 September 2017
Mol Syst Biol. (2017) 13: 947

Introduction

Recent years have seen a rapid development in genome engineering, which, in combination with decreased costs for DNA oligonucleotide synthesis, have made it possible to design and produce pool-generated cell libraries with overwhelming genetic diversity (Wang *et al*, 2009; Dixit *et al*, 2016; Jaitin *et al*, 2016; Peters *et al*, 2016; Garst *et al*, 2017; Otoupal *et al*, 2017). A similarly impressive development in microscopy enables the investigation of complex phenotypes at high temporal resolution and spatial precision in living cells (Liu *et al*, 2015; Balzarotti *et al*, 2017). Biological imaging has benefited greatly from developments in microfluidics which have enabled well-controlled single-cell observations of individual strains over many generations (Wang *et al*, 2010; Uphoff *et al*, 2016; Wallden *et al*, 2016). Despite the rapid technological progress within these areas, there is currently no efficient technique for mapping phenotypes related to intracellular dynamics or localization to their corresponding genotype for pool-generated libraries of genetically different cell strains. Recent work observing multiple bacterial strains on agarose pads allows for sensitive microscopy (Kuwada *et al*, 2015; Shi *et al*, 2017), but the genetic diversity is capped since the strain production and handling is not pooled. On the other end, droplet fluidics allows working with large genetic diversity (Dixit *et al*, 2016) but cannot be used to characterize phenotypes that require sensitive time-lapse imaging.

Here, we present a method that solves the problem by *in situ* genotyping the library of strains after the phenotypes have been studied in time-lapse microscopy. Thus, the genotype of the cell is not known at the time of phenotyping but revealed through the spatial position of the cell after fixation and *in situ* genotyping.

Results

The DuMPLING approach

We refer to our solution of the library phenotyping problem as DuMPLING—**d**ynamic **u**-fluidic **m**icroscopy-based **p**henotyping of a **l**ibrary before ***in situ*** **g**enotyping. DuMPLING is composed of three key components: strain generation, live cell phenotyping, and *in situ* genotyping (schematically outlined in Fig 1). All three components can be made in different ways, but in the current study, we have selected this implementation:

1 *Pool-generated strain library:* We have constructed a library of CRISPRi/dCas9 knockdowns. We generated a recipient strain harboring chromosomal inducibly expressed dCas9 and T7 polymerase. We used Golden Gate assembly to generate a small plasmid-expressed library of sgRNA spacers (to direct the dCas9 chromosomal binding and create knockdowns) and neighboring barcode sequences (for later genetic identification) (Figs 2, EV1, and EV2). Note that in 167 nt, we fit the variable regions (i.e., the barcode sequence and sgRNA spacer sequence), the constant elements between the variable regions and the constant regions on the ends for PCR and assembly (see Supplement for sequence design details). This length of oligo is easily procured from companies, and much larger libraries have been built following this approach with purchased oligo pools (Dixit *et al*, 2016), making it clear that this strategy of library

Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden
*Corresponding author. Tel: +46 18 4714678; E-mail: johan.elf@icm.uu.se
†These authors contributed equally to this work

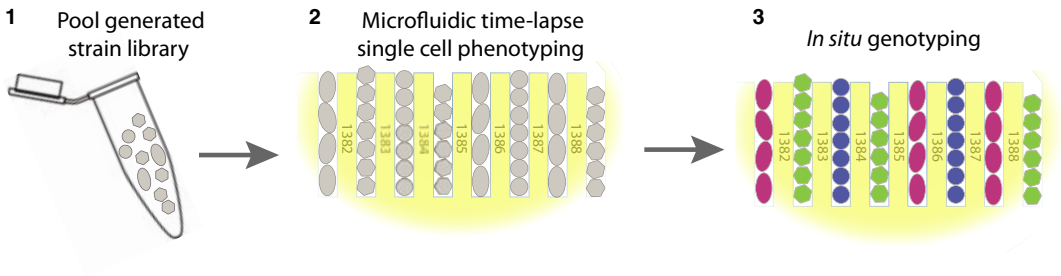

Figure 1. The DuMPLING strategy.
(1) Pooled strain library generation. (2) Live single cell phenotyping using microscopy. (3) Genotypes recovered by *in situ* genotyping.

construction can be extended to a genomewide knockdown library.

2 *Live cell phenotyping in a microfluidic device where each strain occupies a defined position:* The mixed strains are loaded into a microfluidic chip which harbors 4,000 cell channels, sustains continuous exponential growth, and allows single-cell imaging for days (Fig 3A, Movies EV1 and EV2). After only a few generations, all cells in a channel are the progeny of the cell at the back of the channel and thus share the same genotype. The chip design is similar to the mother machine (Wang *et al*, 2010), but we have introduced a 300 nm opening in the back of each cell channel such that media and reagents can be passed over the cells. This redesign facilitates cell loading and is essential for genotyping.

3 *In situ genotyping to identify which strain is in which position*: As mentioned above, each plasmid expresses a unique RNA-based barcode that allows genotype identification. The barcode is expressed from a T7 promoter, and the T7 polymerase is under control of an inducible arabinose promoter. The orthogonal and inducible nature of this system prevents it from interfering with cell physiology during phenotyping. After induction of the barcode RNA expression, the cells are fixed *in situ* with formaldehyde and permeabilized in 70% EtOH before sequential fluorescent *in situ* hybridization (FISH). The individual barcodes are identified by sequential hybridization of fluorescent 37-nt-long oligonucleotides (probes). The multiplexed process of designing and producing the probe library is described in the Materials and Methods section. The templates for probe synthesis are procured in the same array format as the barcoded sgRNA templates. Here, we use probes of two different colors in two sequential rounds of probing, which is sufficient for identifying the three genotypes in this study.

In general, C^N genotypes can be identified where C is the number of colors and N is the number of rounds of probing. Thus, genotyping can straightforwardly be extended to more strains by using more colors or rounds of probing. For example, a recent publication (Shah *et al*, 2016) showed four rounds of single-molecule FISH probing in five colors (i.e., 625 genotypes), and they observed a miss-call rate of ~1%. We would however expect a lower error rate than this as we are imaging ~6 cells of the same genotype, each containing many RNA rather than individual RNA molecules. To demonstrate that it is possible to reprobe many times, we perform $N = 6$ consecutive rounds (Fig 3B) of probing in each position. It is however likely that more rounds are possible without loss of specificity. For example, in a recent study, Chen *et al* were able to successfully probe single RNA molecules 16 rounds (Chen *et al*, 2015).

Proof-of-concept demonstration

To exemplify the use of DuMPLING, we performed targeted knockdowns of different components of the *lac* operon in *Escherichia coli* using a set of sgRNA-expressing plasmids that repressed *lacY*, an unrelated gene or *lacI* (Fig 2A–C). As described above, the plasmids are made from pooled oligos including the sgRNA and its unique barcode. The pooled approach has previously been used to generate libraries of thousands of genotypes (Dixit *et al*, 2016; Jaitin *et al*,

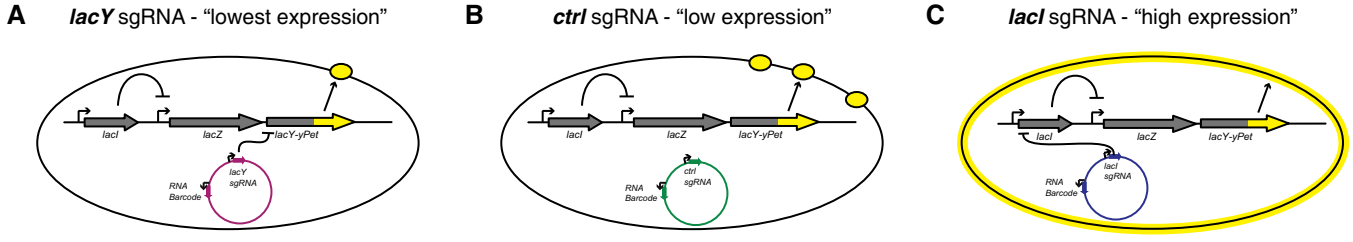

Figure 2. Three strain *lac* operon knockdown library: Repression network for the three different plasmids used.
A *lacY* knockdown (lowest LacY-YPet expression, purple).
B No knockdown (low LacY-YPet expression, green).
C *lacI* knockdown (high LacY-YPet expression, blue).
Data information: Color scheme holds throughout this paper.

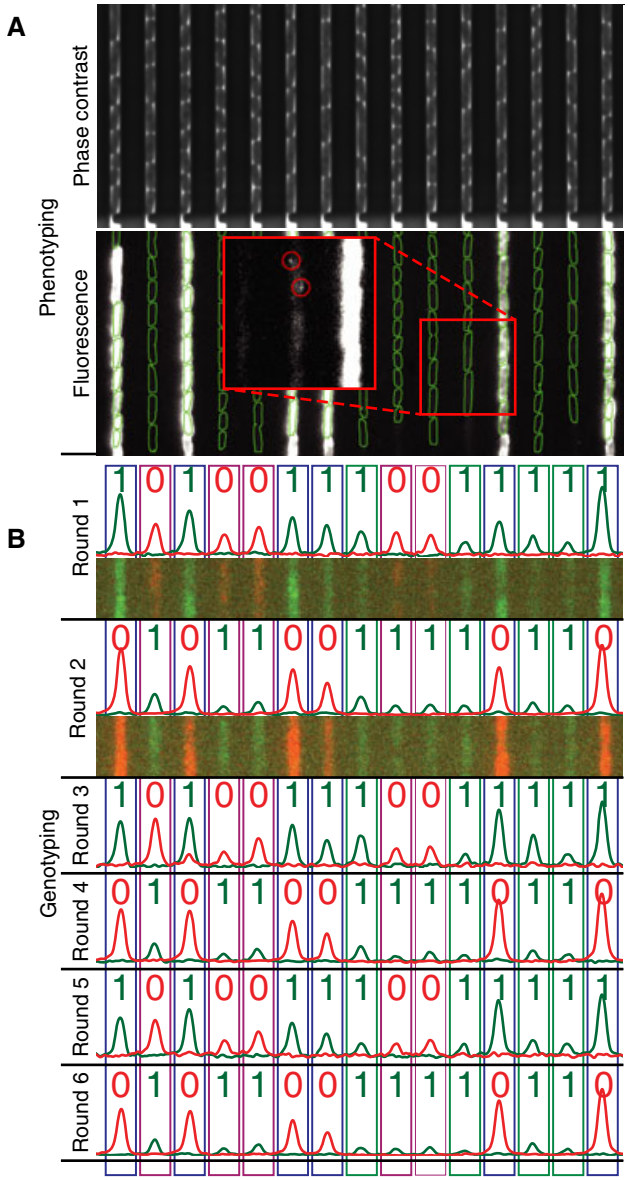

Figure 3. Mapping phenotypes to genotypes.

A Examples of channels and cells in the custom-made microfluidic device which are imaged in both phase contrast (top) and fluorescence microscopy (bottom). Phase contrast is used to segment the cells (green outlines), and single-molecule fluorescence microscopy is used to detect gene expression (red circles in red inset box, which is a blow up of the figure as indicated by the smaller red square and has a change of levels to allow visualization of single molecules) from the *lac* operon.

B *In situ* genotyping with six sequential rounds of FISH probe hybridization and stripping. Cropped images of two cells that are representative of all cells in the trap are shown for the first two rounds (overlay of Cy3 (green) and Cy5 (red) images). The genotype is called by summing the signal in the channel: 0 is assigned for Cy5 (red) and 1 for Cy3 (green). Rectangles indicate assigned genotype (10: lacI knockdown; 01: lacY knockdown; 11: no knockdown).

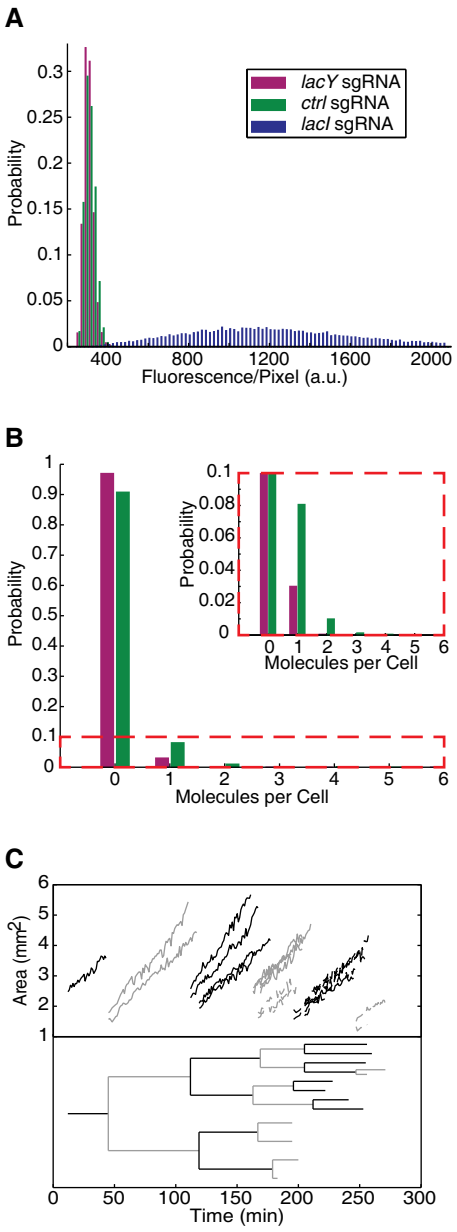

Figure 4. Phenotype data.

A Gene expression categorized by assigned genotype.

B Single-molecule counting of expression from the two low-expression genotypes.

C Top: Growth curves for one cell lineage (from one channel). Dashed lines indicate the end of detection of a branch. Bottom: Corresponding lineage tree.

2016), but here, we limit to three variants to be able to precisely evaluate the accuracy of each step. The mixed plasmids are electroporated into an *E. coli* strain, where dCas9 is expressed from a regulated chromosomal promoter (the promoter is tightly regulated to prevent bias in growth before loading and induction, Fig EV3). Furthermore, the *lacY* gene is fused with the gene for the fluorescent protein YPet to obtain a detectable single-molecule phenotype.

In our experiments, 233 channels are imaged every 60 s using phase contrast and every 13 min using single-molecule-sensitive wide-field fluorescence for a total of 272 min. Phase contrast images are used for cell detection and lineage tracking. Individual

LacY-YPet molecules, detected using wide-field epifluorescence, are overlaid on the phase contrast images to allow assignment of individual molecules to individual cells.

We were able to track a cell lineage over the full time-course of the experiment (six generations) and quantify the growth curves of each member of the family tree (see example in Fig 4C). In addition, the long time course of single-cell/single-molecule microscopy allowed us to reproducibly measure mean expression of less than one YPet molecule per generation and distinguish a $< 3\times$ change at this expression level (compare distribution of single-molecule counts per cell in Figs 4B and EV4). This type of phenotyping is not possible in most other settings (e.g., flow cytometry) and would not scale to hundreds of strains in those where it is possible (e.g., agarose pads mounted on a microscope).

While the phenotypic difference between the two low-expression strains can only be resolved with extensive single-molecule time-lapse imaging, we also included the *lacI* knockdown phenotype, which is trivial to identify, to test for correct genotype to phenotype assignments. All 74 channels with cells that express a high level of LacY-YPet (Fig 3A) have been correctly found to express the barcode RNA associated with the sgRNA against *lacI* (blue boxes in Fig 3B and blue bars in Fig 4A), and all channels with cells with the barcode RNA associated with the sgRNA against *lacI* express high levels of LacY-YPet. The observed sensitivity and specificity for identifying the genotype in this experiment is therefore 100%. If we also consider the limited sample size and the redundant genotyping as independent, the sensitivity is > 97.5% and the specificity > 99% (see Materials and Methods section for details).

Discussion

This paper describes a proof-of-principle application of the DuMPLING concept, that is, the possibility to use advanced microscopy to phenotype a pool-generated library of live cells and then genotype *in situ*. The advantage of our method compared to the state of the art is the combination of pooled handling of library generation and characterization of complex phenotypes based on dynamic changes in single cells. We have used a microfluidic device to both phenotype the cells in a constant growth environment for an extended period of time and perform the subsequent genotyping.

We note that each of the components (strain library generation, phenotyping, and genotyping) can be performed in different ways depending on the specific question. For example, one can make pooled dCas9 libraries based on plasmids harboring both a genotype-identifying barcode and a sgRNA gene (Dixit *et al*, 2016; Jaitin *et al*, 2016; Peters *et al*, 2016; Garst *et al*, 2017; Otoupal *et al*, 2017) for labeling genetic loci (Chen *et al*, 2013) or knocking down/activating genes throughout the chromosome. Alternatively, pooled chromosomal libraries with variants of promoters, ribosome binding sites (RBS), or coding sequences (Wang *et al*, 2009; Keren *et al*, 2016) can be made. Furthermore, it is in general not necessary to introduce the barcode in direct proximity to the genetic alterations as long as the barcode can be connected to the genotype in some other way than through the oligo synthesis. For example, long sequence reads can connect random barcodes to the genetic alteration that causes a phenotype.

Similarly, sensitive single-cell time-lapse imaging can be used to characterize a bewildering diversity of cell phenotypes than are not accessible with snapshot measurement as obtained in FACS or droplet fluidics (Norman *et al*, 2013; Hammar *et al*, 2014; Taheri-Araghi *et al*, 2015; Potvin-Trottier *et al*, 2016; Wallden *et al*, 2016). Depending on the cell types and the experiment, it may also be more convenient to use an open culture dish instead of the fluidic device.

Also, the method for identifying the barcode can be implemented in different ways such as *in situ* sequencing (Ke *et al*, 2013; Lee *et al*, 2014). One advantage of direct *in situ* sequencing is that the genotype may be identified directly without the use of a barcode.

In short, while we have presented a CRISP-FISH-DuMPLING, the DuMPLING can have many other fillings.

Materials and Methods

Reagents and tools table

Reagent/Resource	Reference or source	Identifier or catalog number
Experimental models		
C57BL/6J (*M. musculus*)	Jackson Lab	B6.129P2Gpr37tm1Dgen/J
DH5alpha (*E. coli*)	ThermoFisher	Cat # 18265017
BW29655 *(E. coli)*	Zhou *et al*., 2003	N/A
NIH 3T3 cells (*M. musculus*)	ATCC	Cat #
Liver patient biopsies	Heidelberg University Hospital	N/A
Ap-GAL4 (*D. melanogaster* strain)	Bloomington Drosophila Stock Center	BDSC:3041; FLYB:FBti0002785
Recombinant DNA		
pCMV-BE3	Addgene	Cat #73021
pBRAFV600E (*H. sapiens*)	This study	N/A
pBRAF (*M. musculus*)	This study	N/A

Continued

Reagent/Resource	Reference or source	Identifier or catalog number
pEYFP-Myosin (*D. melanogaster*)	J. James lab, Smith *et al*, 2017	N/A
pSR43.6 (CcaSR)	Schmidl *et al*., 2014	N/A
Antibodies		
Rabbit anti-H3	Abcam	Cat#ab1791
Goat anti-Cy3	Cedarlane	Cat#111-165-003
Mouse anti a-Tubulin monoclonal antibody (clone DM1A)	Sigma Aldrich	Cat #T9026
Rabbit polyclonal anti-Nanog antibody	This study	N/A
Oligonucleotides and other sequence-based reagents		
Cloning oligos	This study	Table 1
PCR primers	This study	Table EV3
siRNA sequences	This study	Table EV5
shRNA sequences		
Genotyping primers	This study	Table 1
Chemicals, enzymes and other reagents		
Kanamycin	Sigma-Aldrich	Cat. #K0879
T7 Endonuclease I	New England Biolabs	Cat # M0302S
Protease Inhibitor Cocktail	Roche	Cat# 04693159001
MEK1/2 inhibitor U0126	Cell Signaling	Cat # #9903
Vemurafenib	MedChem Express	Cat#: HY-12057
Insulin	Sigma	Cat #:I2643
Software		
Cytoscape v3.4.0	http://www.cytoscape.org Shannon *et al*, 2003	
Perseus	Tyanova *et al*, 2016 http://www.coxdocs.org/doku.php?id=perseus:start	
CellProfiler	Carpenter *et al*, 2006 http://cellprofiler.org	
GECKO method	This study, https://github.com/SysBioChalmers/GECKO/releases/tag/v1.0	
DeepLoc	Kraus et al, 2017 https://github.com/okraus/DeepLoc	
Other		
Eclipse Ti-E Inverted TIRF Microscope	Nikon	MEA53100
Opera LX Spinning Disk Microscope	Perkin Elmer	
Plate reader	Tecan	Infinite M200 Pro
Illumina NextSeq 500	Illumina	
CellCarrier 384-well glass-bottom imaging plates	Perkin Elmer	Cat # 6007550

Methods and Protocols

Full-length human bait generation

Each human GPCR was amplified by PCR and inserted by homologous recombination (Chen *et al.,* 1992) in yeast into either of the two bait vectors pCCW-STE or pTMBV (Dualsystems Biotech). The primers used for the pCCW vector are 5′-CCTTTAATTAAGG CCGCCTCGGCCATCTGCAGG-3′ (forward) and 5-CGACATGGTCGA CGGTATCGATAAGCTTGATATCAGCAGTGAGTCATTTGTACTAC-3′ (reverse). The primers used for the pTMBV4 vector are 5′- CCAG TGGCTGCAGGGCCGCCTCGGCCAAAGGCCTCCATGG-3′ (forward) and 5- ATGTCGGGGGGGATCCCTCCAGATCAACAAAGATTG-3′ (reverse). In MYTH bait vectors, the GPCRs were fused N-terminally to the yeast mating factor alpha signal sequence to target full-length non-yeast membrane proteins to the membrane (King *et al.,* 1990). At the C-terminus, the GPCR was fused in-frame with the MYTH tag

consisting of a C-terminal ubiquitin (Cub) moiety and LexA-VP16 transcription factor (TF) (Fashena, SJ, Serebriiskii, IG, Gomelis, 2000; Fields & Song, 1989).

Bait validation

The resulting MYTH bait constructs were tested as previously described (Snider *et al*, 2013, 2010). Briefly, the baits were transformed (Gietz & Woods, 2006) into either of the yeast reporter strains THY.AP4 or NMY51. The correct localization of modified baits to the membrane was confirmed by immunofluorescence using (rabbit) anti-VP16 (Sigma Cat# V4388) (1/200); secondary (goat) anti-(rabbit) Cy3 (Cedarlane Cat#111-165-003) (1/500)). Test MYTH was carried out with control interacting (NubI) preys to confirm functionality in MYTH, and with non-interacting (NubG) preys to verify that baits do not self-activate in the absence of interacting prey (Snider *et al.,* 2010).

Functionality of select GPCR-Cub-TF baits (Pausch, 1997) was confirmed (Dowell & Brown, 2009) in either wild-type THY.AP4 or the same strain expressing a given GPCR-Cub-TF fusion. Cells were diluted from an overnight culture to an OD600 of 0.0625 in minimum SD or SD-Leu media, respectively. The various concentrations of drug, Salmeterol (agonist for ADRB2) or morphine (agonist for OPRM1), were added to a final concentration of 200 μM. The growth rate was monitored by measuring the OD600 every 15 minutes for 24 hours by TECAN Sunrise plate reader.

Confirmation of known GPCR interactions by MYTH

Known GPCR interacting partners were identified from the Integrated Interactions Database (IID) (Ritter & Hall, 2009; Kotlyar *et al*, 2016; Chung *et al*, 2013). Gateway compatible ORFs were obtained from the Human ORFeome Collection version 8.1 (Yang *et al.,* 2011) and used, via the Gateway system (Life Technologies), to generate either N-terminally tagged preys in pGPR3N (Dualsystems Biotech) or C-terminally tagged preys in pGLigand (created in-house, Stagljar lab) depending on which end is available for tagging. All bait prey interaction tests were carried out using MYTH as previously described (Snider *et al*, 2010) in the NMY51 yeast reporter strain. Note that prior to use in interaction tests with GPCR baits all preys were tested for promiscuity by use of an artificial bait construct that consists of the single-pass transmembrane domain of human T-cell surface glycoprotein CD4 and the Cub-TF tag (Snider *et al.,* 2010) and by use of the yeast protein RGT2.

Membrane yeast two hybrid (MYTH) screens

1) Transform bait containing yeast in duplicate with the human fetal brain DUALmembrane cDNA library in the N_{ub}G-x orientation (DualSystems Biotech) as previously described (Snider *et al*, 2010).
2) Plate onto synthetic dropout minus tryptophan, leucine, adenine and histidine (SD-Trp-Leu-Ade-His) plates with various amounts of 3-Amino-1,2,4-triazole (3-AT) as assessed by the N_{ub}G/I control test for each individual bait.
3) Pick transformants and spot onto SD-Trp-Leu-Ade-His plates containing 3-AT and X-Gal dissolved in *N,N*-dimethyl formamide.
4) Use blue colonies, expressing putative interacting preys, to inoculate overnight liquid cultures (SD-Trp).
5) Extract plasmid DNA from these cultures and use it to transform *E.coli*, DH5alpha strain for amplification.
6) Extract the plasmid DNA once more. Send for sequencing and use in the bait dependency test to rule out spurious interactors, as described previously (Snider *et al*, 2010).

Filtering interactions

To reduce the number of false positives, we eliminated detected interactions involving preys that carry out signal peptide processing (GO:0006465) and ribosomal contaminants (Glatter *et al.,* 2009). We identified these preys using Gene Ontology (GO) (Ashburner *et al.,* 2000) annotations from the UniProt-GO Annotation database (Dimmer *et al.,* 2012; Matthews *et al.,* 2009), downloaded through the EMBL-EBI QuickGO browser (Binns *et al.*, 2009) (http://www.ebi.ac.uk/QuickGO/GTerm?id = GO:0006465#term = annotation), on Sept. 10, 2016.

Identifying previously known interactions

Overlap between detected interactions and interactions already reported in previous studies was identified using the IID database (Kotlyar *et al*, 2016) ver. 2016-03 (http://ophid.utoronto.ca/iid).

Annotating interacting proteins: Membrane localization

Baits and preys localized to the plasma membrane were identified using GO annotations from the UniProt-GO Annotation database (Dimmer *et al.,* 2012), obtained through the EMBL-EBI QuickGO browser (Binns *et al.*, 2009) (http://www.ebi.ac.uk/QuickGO/GTerm?id = GO:0006465#term = annotation) on Aug. 31, 2016.

Process annotations and enrichment analysis

Baits and preys were annotated with GO Slim process terms from the *goslim_generic* set (http://www.ebi.ac.uk/QuickGO/GMultiTerm#tab = choose-terms) (Table EV3). We downloaded annotations on Aug. 31, 2016.

Pathway annotations

Pathway annotations for baits and preys, as well as pathway enrichment analysis, were performed using the pathDIP database (Rahmati *et al*, 2016) ver. 2.5 (http://ophid.utotoronto.ca/pathDIP), using the setting 'Extended pathway associations' with default parameters. *P*-values were FDR-corrected using the Benjamini-Hochberg method.

Disease annotations and enrichment analysis

Disease annotations for baits and preys were downloaded from the DisGeNET database (Piñero *et al*, 2015) v4.0, on Aug. 31, 2016. Disease enrichment of preys was assessed by calculating hypergeometric *P*-values (using the human genome as the background population), and correcting for multiple testing using the Benjamini-Hochberg method.

Molecular function and biological process annotations and enrichment analysis

Molecular function and biological process Gene Ontology annotations (161130) were downloaded from Gene Ontology Consortium (Gene Ontology Consortium, 2015). Enrichment of preys for molecular functions was calculated using the topGO library version 2.24.0 in R version 3.3.1 (Alexa & Rahnenfuhrer, 2016). A topGOdata object was created with nodeSize = 10 and the runTest function was used with the default algorithm (weight01) and statistic = fisher. *P*-values were adjusted for multiple testing using the Benjamini-

Hochberg method. Enrichment of preys for biological processes was calculated the same way.

Domain annotation and enrichment analysis

InterPro domain annotations were obtained from UniProt release 2016_11 (Mitchell *et al*, 2015; UniProt Consortium, 2015). Domain enrichment of preys was assessed by calculating hypergeometric *P*-values (using the human proteome as the background population), and correcting for multiple testing using the Benjamini-Hochberg method.

Domain pairs enriched among interacting bait-prey pairs was identified in 2 steps. First, sets of co-occurring domains were identified for baits; each set comprised domains that always occurred together on baits. Similarly, sets of co-occurring domains were identified on preys. Domains that did not always co-occur with others were considered domain sets of length 1. Enrichment was subsequently calculated for pairs of domain sets - one set on baits and the other on preys. Domain sets were identified for three reasons: (1) to avoid redundant results from different domains representing the same proteins, (2) to avoid excessive multiple testing penalties from non-independent tests, and (3) for easier interpretation of results, since a domain set clarifies that enrichment analysis cannot distinguish between domains within the set. After domain sets were identified, *P*-values were calculated for domain set pairs using hypergeometric probability with the following parameters: N = the number of possible interactions involving baits (number of baits x size of human proteome), M = the number of detected interactions, n = the number of possible pairings between the bait domain set and the prey domain set (number of baits with domain set x number of human proteins with prey domain set), and m = number of interacting bait-prey pairs with corresponding domain sets. Adjusted *P*-values were calculated using the Benjamini-Hochberg method.

Drug target enrichment and drug category enrichment

Drug targets and drug therapeutic categories were downloaded from DrugBank version 5 (Wishart *et al*, 2006). We calculated target enrichment among GPCR baits as a hypergeometric *P*-value, using the following parameters: the number of human protein-coding genes in the HGNC database (Gray *et al*, 2015) (n = 19,008), the number drug targets in DrugBank (n = 4,333), the number of baits (n = 48), and the number of baits that are drug targets (n = 28).

Enrichment of therapeutic categories among baits and preys was calculated as hypergeometric *P*-values using the following parameters: the number of human protein-coding genes in the HGNC database (Gray *et al*, 2015) (n = 19,008), the number of targets in a therapeutic category, the number of baits and preys (n = 686), and the number of baits and preys that are targets in the category. We calculated *Q*-values (*P*-values adjusted for multiple testing) using the Benjamini-Hochberg method.

Drugs sales and prescription numbers were obtained from Medscape (100 Best-Selling, Most Prescribed Branded Drugs Through March. Medscape. May 06, 2015).

PPI predictions

Predictions were obtained using the FpClass algorithm (Kotlyar *et al*, 2015): a probabilistic method that integrates diverse PPI evidence including compatibility of protein domains, gene co-expression, and functional similarity, as well as other methods integrated in IID (Kotlyar *et al*, 2016). Resulting networks were visualized in NAViGaTOR 3.0 (http://ophid.utoronto.ca/navigator; (Brown *et al*, 2009)).

Confirmation of interactions by Co-immunoprecipitation

Approach 1—Endogenous baits and transiently transfected FLAG-tagged preys

Maintain 293T cells in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS, 100 U penicillin and 100 μg/ml streptomycin (Fisher Scientific, cat# SV30010) and split at 80% confluence. To co-immunoprecipitate GPCRs with their preys, transiently transfect plasmids encoding FLAG-tagged preys in the 293T cells. Detect their interaction with GPCR using Western blotting with anti-GCPR antibodies.

The procedure is as follows:

- Plate 293T cells at 40% confluence overnight.
  - On the following morning, transfect cells using calcium phosphate [$Ca_3(PO_4)_2$] kit ProFection from Promega (cat# E1200) following manufacturer's instructions.
  - Add 70 μg of plasmid DNA to CaCl2 and water and add the mixture to HEPES-buffered saline while vortexing.
- Incubate the mixture at room temperature for 30 minutes.
- Prior to adding to cells, vortex the mixture again.
  - 24 hours post-transfection, harvest 2 X 150 mm dishes of 293T cells/plasmid and wash the cells with ice-cold PBS.
  - After that, cross-link the cells with 0.5 mM DSP at room temperature for 30 mins followed by quenching excessive DSP with a buffer containing 0.1 M Tris-HCl, pH 7.5 and EDTA 2 mM.
  - Centrifuge detached cells at 1400 rpm for 10 minutes at 4°C.
  - Lyse the cell pellet in RIPA buffer containing 1X protease inhibitor cocktail (Sigma Aldrich, cat# P2714) on ice for 30 minutes with occasional agitation. To aid lysis, pass the cells through a 21G needle 10X.
- Clear the lysate by centrifugation at 13,000 RPM for 15 minutes at 4 °C.
- Adjust a volume of cell lysate containing 10 mg protein to 1 ml with RIPA containing 1X protease inhibitor cocktail and add 3 μg of each anti-GPCR receptor antibody.
  - Rotate the tube for 1 hour at 4 °C followed by addition of 100 μl of μMACS protein-G magnetic microbeads (Miltenyi, cat# 130-071-101) with continued rotation for additional 4 hours at 4 °C.
  - Equilibrate μMACS columns (Miltenyi, cat# 130-092-444) with RIPA 1X protease inhibitor complex.
  - Pass the microbeads suspension through the columns and wash the retained microbeads 3X with 800 μl of RIPA 0.1% of detergents and 1X protease inhibitor cocktail followed by another 2X washes with 500 μl detergent-free RIPA containing 1X protease inhibitor cocktail only.
- Release the proteins bound to the microbeads by adding 25 μl Laemmli loading buffer at 95 °C 2X.
  - Analyze eluates using SDS-PAGE and visualize using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher, cat# 34094).

Approach 2—Endogenous baits and preys

- Harvest ten 150 mm dishes of HEK-293 cells and centrifuge at 400 *x g* for 10 min.
- Resuspend the cell pellet in 15 ml phosphate-buffered saline

(PBS) and mixed with an equal volume of cross-linking reagent (1 mM dithiobis-succinimidyl propionate prepared in PBS).
- After 30 min incubation, pellet the cross-linked cells by centrifugation at 400 *x g*, and lyse in IPLB (immunoprecipitation lysis buffer containing 1% digitonin and 1X protease inhibitor cocktail) for 30 min.
- Centrifuge the lysates at 16,000 *x g* for 15 min at 4 °C.
- Adjust the cell lysate containing ~10 mg of protein to 1 ml with IPLB (containing 1% digitonin and 1X protease inhibitor cocktail) and add 3 μg of antibody specific to the target protein to the mixture.
- Incubate the samples with 100 μl of μMACS protein-G magnetic beads followed by 5 h gentle rotation at 4 °C.
- Pass the bead suspension through the μMACS columns (equilibrated with IPLB containing 1% digitonin and 1X protease inhibitor), and wash the retained beads 3 times with 800 μl of IPLB (0.1% digitonin and 1X protease inhibitor) followed by another 2 washes with 500 μl IPLB (1X protease inhibitor only).
- Elute co-purifying protein that bound to the beads by the addition of 25 μl Laemmli loading buffer at 95 °C, and analyze by SDS-PAGE and immunoblotting using protein specific antibody.

Antibodies used in co-immunoprecipitation experiments

Santa Cruz: OPRL1 (sc-15309), TSHR (sc-13936), OPRM1 (sc-15310), AGTR1 (sc-1173-G), PTAFR (sc-20732), C5L2 (sc-368573), HRH (sc-20633), CHRM5 (sc-9110), OXTR (sc-33209).

Abcam: ADRB2 (ab36956), HNRPK (ab52600), F2RL (ab124227), TTYH1 (ab57582), PRNP (ab52604), MGLL (ab24701), ATP2A2 (ab2861), FA2H (ab54615), HSPA1B (ab79852).

Cell Signaling: GABBR1 (3835).

ProteinTech: GPR37 (14820-1-AP), FZD7 (16974-1-AP).

Confirmation of interactions by BRET

To confirm select interactions using BRET as an orthogonal validation assay, GPCR interactors identified in MYTH assays were fused to GFP2, a blue-shifted variant of GFP, to act as BRET acceptor, and GPCR receptors to RLucII, a brighter *Renilla* luciferase mutant, to act as donor, then plotted as increasing BRET levels compared to GFP/Rluc, as previously described (Mercier *et al*, 2002; Breton *et al*, 2010; Loening *et al*, 2006).

5-HT4d experiments

The cDNAs encoding human GPR37 and GPRIN2 were purchased from UMR cDNA Resource Center. The 5-HT4d-*R*luc, 5-HT4d-YFP and HA-CCR5 constructs have been described elsewhere (Berthouze *et al.,* 2005; Tadagaki *et al.,* 2012). An N-terminally 6xMyc tagged version of GPRIN2 and GPR37 and C-terminally YFP tagged GPR37-YFP and GPRIN2-YFP fusion proteins were obtained by PCR using the Phusion High-Fidelity DNA Polymerase (Finnzymes). All constructs were inserted in the pcDNA3.1 expression vector and verified by sequencing. The C-terminally deleted GPRIN2Cter construct was obtained by mutagenesis by introducing a stop codon resulting in a truncated protein of 149 amino acids.

Co-immunoprecipitation

HEK-293 cells transiently transfected with 5-HT4d-YFP and myc-GPRIN2 or GPR37 were analysed in the presence and absence of 1 μM 5-HT for 15 minutes and processed for immunoprecipitation using a monoclonal anti-GFP antibody. Crude extracts and immunoprecipitates were analysed by SDS-PAGE and immunoblotted using rabbit anti-GFP or anti-myc antibodies.

BRET

BRET donor saturation curves were performed in HEK-293 cells by co-transfecting a fixed amount of 5-HT4d-*R*luc and increasing amounts of 5-HT4d -YFP, GPR37-YFP and GPRIN2-YFP as described previously (Maurice *et al.,* 2010)

Fluorescence microscopy

HELA cells expressing 5-HT4d-YFP and Myc-GPR37 or Myc-GPRIN2 were fixed, permeabilized with 0.2% Triton X-100, nuclei stained with DAPI (blue) and incubated with monoclonal anti-Myc antibody (Sigma, St Louis, MO) (2 mg/ ml) and subsequently with a Cy3-coupled secondary antibody. GFP, Cy3 and DAPI labeling was observed by confocal microscopy.

Signaling assays

ERK1/2 activation and cyclic AMP levels were determined in HEK-293 cells as described previously (Guillaume *et al.,* 2008).

ADORA2A experiments

The cDNA encoding the human GPR37 (Unigene ID: Hs.725956; Source BioScience, Nottingham, U.K.) was amplified and subcloned into the HindIII/EcoRI restriction sites of the pEYFP vector (Invitrogen, Carlsbad, CA, USA) using the iProof High-Fidelity DNA polymerase (Bio-Rad, Hercules, CA, USA) and the following primers: FGPR37 (5′-CGCAAGCTTATGCGAGCCCCGG-3′) and RGPRYFP (5′-CGCGAATTCCGCAATGAGTTCCG-3′). GPR37 was also subcloned in the HindIII/KpnI restriction sites of the p*R*luc-N1 vector (Perkin–Elmer, Waltham, MA, USA) using the following primers FGP37 and RGP*R*Luc (5′-CGCGGTACCGCGCAATGAGTTCCG-3′).

The constructs for the human adenosine A2A receptor (namely, ADORA2A-YFP and ADORA2A-*R*luc) were obtained as previously described (Gandia *et al.,* 2008) and ADORA2A-CFP was obtained by subcloning the adenosine receptor from ADORA2A-YFP into the pECFP-N1 plasmid.

A homemade rabbit anti-GPR37 polyclonal antibody (Lopes *et al*, 2015) was used. Other antibodies used were rabbit anti-A_{2A}R (Ciruela et al., 2004), mouse anti-A_{2A}R (05-717, Millipore, Temecula, CA, USA), rabbit anti-Flag (F7425, Sigma) and rabbit anti-A_1R (PA1-041A, Affinity BioReagents, Golden, CO, USA).

C57BL/6J wild type and GPR37-/- mice with a C57BL/6J genetic background (Strain Name: B6.129P2-Gpr37tm1Dgen/J; The Jackson Laboratory, Bar Harbor, ME, USA) were used. Mice were housed in standard cages with ad libitum access to food and water, and maintained under controlled standard conditions (12 h dark/light cycle starting at 7:30 AM, 22°C temperature and 66% humidity). The University of Barcelona Committee on Animal Use and Care approved the protocol and the animals were housed and tested in compliance with the guidelines described in the Guide for the Care and Use of Laboratory Animals (Clark et al., 1997) and following the European Community, law 86/609/CCE.

Immunocytochemistry

HEK-293 cells were transiently transfected with ADORA2A-CFP, GPR37-YFP or ADORA2A-CFP plus GPR37-YFP using the

TransFectin Lipid Reagent (Bio-Rad) and following the instructions provided by the manufacturer. The cells were analyzed by confocal microscopy 48 h after transfection. Superimposition of images (merge) reveals co-distribution of ADORA2A-CFP and GPR37-YFP in yellow and DAPI-stained nuclei in blue. Scale bar: 10 μm.

Co-immunoprecipitation

- Obtain membrane extracts from HEK-293 cells and C57BL/6J mouse striatum as described previously (Burgueño *et al*, 2003).
- Solubilize the membranes in ice cold radioimmunoassay (RIPA) buffer (150 mM NaCl, 1% NP-40, 50 mm Tris, 0.5% sodium deoxycholate, and 0.1% SDS, pH 8.0) for 30 min on ice in the presence of protease inhibitor (Protease Inhibitor Cocktail Set III, Millipore, Temecula, CA, USA).
- Centrifuge the solubilized membrane extract at 13,000 × g for 30 min and incubate the supernatant overnight with constant rotation at 4 °C with the indicated antibody.
- Add 50 μl of a suspension of Protein A-Agarose (Sigma) or True-Blot anti-rabbit Ig IP beads (eBioscience, San Diego, CA) and incubate for another 2 hours.
- Wash the beads with ice-cold RIPA buffer, dissociate immune complexes, transfer to polyvinylidene difluoride membranes and probe with the indicated primary antibodies followed by horseradish-peroxidase (HRP)-conjugated secondary antibodies.
- Detect the immunoreactive bands using Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific) and visualize in a LAS-3000 (FujiFilm Life Science).

BRET saturation eperiments

- Transiently transfect HEK-293 cells with a constant amount of cDNA encoding the *R*luc constructs and increasing amounts of YFP tagged proteins.
- Rapidly wash twice in PBS, detach and resuspend in Hank's balanced salt solution (HBSS) buffer (137 mM NaCl, 5.4 mM KCl, 0.25 mM Na2HPO4, 0.44 mM KH2PO4, 1.3 mM CaCl2, 1.0 mM MgSO4, 4.2 mM NaHCO3, pH 7.4), containing 10 mM glucose.
- Process for BRET determinations using a POLARstar Optima plate-reader (BMG Labtech, Durham, NC, USA) (Ciruela *et al.*, 2015) or Mithras plate reader (Berthold Technologies) (Cecon *et al*, 2015).

Cell surface expression

HEK-293 cells were transiently transfected with the cDNA encoding ADORA2A, ADORA1, GPR37-YFP, ADORA2A plus GPR37-YFP or ADORA1 plus GPR37-YFP. Cell surface labelling was performed by biotinylation experiments (Burgueño *et al*, 2003). Crude extracts and biotinylated proteins were subsequently analyzed by SDS-PAGE and immunoblotted using a rabbit anti-GPR37 antibody (1/2000), a rabbit anti-$A_{2A}R$ antibody (1/2000) or a rabbit anti-A_1R antibody (1/2000). The primary bound antibody was detected as described before.

Catalepsy score

Catalepsy behavior was induced by the D_2R antagonist haloperidol (1.5 mg/kg, i.p.), as previously described (Chen *et al*, 2001). Mice used in the catalepsy test were two months old males. The animals were randomly distributed among the experimental groups. Fifteen minutes before animals were administered either saline or SCH58261 (1 mg/kg, i.p.), an $A_{2A}R$ antagonist. The cataleptic response was measured as the duration of an abnormal upright posture in which the forepaws of the mouse were placed on a horizontal wooden bar (0.6 cm of diameter) at 4.5 cm high from the floor. The latency to move at least one of the two forepaws was recorded 2 h after Haloperidol administration. The test was carried out by an experimenter who was blind to the identity of treatments and the cataleptic time latency was automatically recorded and counted by an independent researcher. A cut-off time of 180 seconds was imposed. Catalepsy testing was performed under dim (16 lux) light conditions. The sample size was initially set as five determinations per experimental condition. Subsequently, the statistical power was calculated using the IBM SPSS Statistics (version 24) software. Accordingly, the sample size was then designed to achieve a minimum of 80% statistical power.

Data availability

The code used for analyzing the data and generating images is provided as Code EV1. Raw images can be downloaded from BioStudies https://www.ebi.ac.uk/biostudies/ (accession code: S-BSST37).

Expanded View for this article is available online.

Acknowledgements

E. Gullberg is gratefully acknowledged for providing the DA25184 strain that carries the spectinomycin resistance cassette and Prune Leroy for providing the LacY-YPet strain. The authors would like to thank George Church and Mats Nilsson for helpful discussions. In addition, the authors would like to acknowledge funding from the Knut and Alice Wallenberg foundation, the Swedish Research Council and the European Research Council.

Author contributions

JE conceived the concept and coordinated the project. MJL developed the phenotyping and genotyping protocols and carried out the corresponding experiments. DC designed the strains and made them. JL and MJL developed the probe synthesis method. ÖB developed the microfluidic device. DF and MJL developed the microscopy and analysis methods. JE, MJL, DF, DC, and JL wrote the manuscript. However, the authors worked closely on the whole project and made substantial contributions in each other's main areas.

Conflict of interest

Concepts related to this work are described in the patent application PCT/SE2015/050227.

References

Baltekin Ö, Boucharin A, Tano E, Andersson DI, Elf J (2017) Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proc Natl Acad Sci USA* 114: 9170 – 9175

Balzarotti F, Eilers Y, Gwosch KC, Gynnå AH, Westphal V, Stefani FD, Elf J, Hell SW (2017) Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* 355: 606 – 612

Beliveau BJ, Joyce EF, Apostolopoulos N, Yilmaz F, Fonseka CY, McCole RB, Chang Y, Li JB, Senaratne TN, Williams BR, Rouillard J-M, Wu C-T (2012) Versatile design and synthesis platform for visualizing genomes with oligopaint FISH probes. *Proc Natl Acad Sci USA* 109: 21301 – 21306

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B (2013) Dynamic imaging of

genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155: 1479 – 1491

Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X (2015) RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348: aaa6090

Clopper CJ, Pearson ES (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 26: 404 – 413

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97: 6640 – 6645

Delebecque CJ, Lindner AB, Silver PA, Aldaye FA (2011) Organization of intracellular reactions with rationally designed RNA assemblies. *Science* 333: 470 – 474

Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, Marjanovic ND, Dionne D, Burks T, Raychowdhury R, Adamson B, Norman TM, Lander ES, Weissman JS, Friedman N, Regev A (2016) Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 167: 1853 – 1866.e17

Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N (2010) Computer control of microscopes using μManager. *Curr Protoc Mol Biol* 92: 14.20.1 – 14.20.17

Engler C, Marillonnet S (2013) Combinatorial DNA assembly using golden gate cloning. *Methods Mol Biol* 1073: 141 – 156

Espah Borujeni A, Channarasappa AS, Salis HM (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res* 42: 2646 – 2659

Garst AD, Bassalo MC, Pines G, Lynch SA, Halweg-Edwards AL, Liu R, Liang L, Wang Z, Zeitoun R, Alexander WG, Gill RT (2017) Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat Biotechnol* 35: 48 – 55

Hammar P, Walldén M, Fange D, Persson F, Baltekin O, Ullman G, Leroy P, Elf J (2014) Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nat Genet* 46: 405 – 408

Imburgio D, Rong M, Ma K, McAllister WT (2000) Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry* 39: 10419 – 10430

Jaitin DA, Weiner A, Yofe I, Lara-Astiaso D, Keren-Shaul H, David E, Salame TM, Tanay A, van Oudenaarden A, Amit I (2016) Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* 167: 1883 – 1896.e15

Ke R, Mignardi M, Pacureanu A, Svedlund J, Botling J, Wählby C, Nilsson M (2013) *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* 10: 857 – 860

Keren L, Hausser J, Lotan-Pompan M , IV S, Alisar H, Kaminski S, Weinberger A, Alon U, Milo R, Segal E (2016) Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell* 166: 1282 – 1294.e18

Kuwada NJ, Traxler B, Wiggins PA (2015) Genome-scale quantitative characterization of bacterial protein localization dynamics throughout the cell cycle. *Mol Microbiol* 95: 64 – 79

Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, Terry R, Jeanty SSF, Li C, Amamoto R, Peters DT, Turczyk BM, Marblestone AH, Inverso SA, Bernard A, Mali P, Rios X, Aach J, Church GM (2014) Highly multiplexed subcellular RNA sequencing *in situ*. *Science* 343: 1360 – 1363

Link AJ, Phillips D, Church GM (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* 179: 6228 – 6237

Liu Z, Lavis LD, Betzig E (2015) Imaging live-cell dynamics and structure at the single-molecule level. *Mol Cell* 58: 644 – 659

Loy G, Zelinsky A (2003) Fast radial symmetry for detecting points of interest. *IEEE Trans Pattern Anal Mach Intell* 25: 959 – 973

Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25: 1203 – 1210

Magnusson KEG, Jalden J, Gilbert PM, Blau HM (2015) Global linking of cell tracks using the viterbi algorithm. *IEEE Trans Med Imaging* 34: 911 – 929

Norman TM, Lord ND, Paulsson J, Losick R (2013) Memory and modularity in cell-fate decision making. *Nature* 503: 481 – 486

Otoupal PB, Erickson KE, Escalas-Bordoy A, Chatterjee A (2017) CRISPR perturbation of gene expression alters bacterial fitness under stress and reveals underlying epistatic constraints. *ACS Synth Biol* 6: 94 – 107

Peters JM, Colavin A, Shi H, Czarny TL, Larson MH, Wong S, Hawkins JS, Lu CHS, Koo B-M, Marta E, Shiver AL, Whitehead EH, Weissman JS, Brown ED, Qi LS, Huang KC, Gross CA (2016) A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* 165: 1493 – 1506

Potvin-Trottier L, Lord ND, Vinnicombe G, Paulsson J (2016) Synchronous long-term oscillations in a synthetic gene circuit. *Nature* 538: 514 – 517

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152: 1173 – 1183

Ranefall P, Sadanandan SK, Wählby C (2016) Fast adaptive local thresholding based on ellipse fit. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 205 – 208.

Rogers JK, Guzman CD, Taylor ND, Raman S, Anderson K, Church GM (2015) Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res* 43: 7648 – 7660

Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27: 946 – 950

Shah S, Lubeck E, Zhou W, Cai L (2016) *In situ* transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* 92: 342 – 357

Shao K, Ding W, Wang F, Li H, Ma D, Wang H (2011) Emulsion PCR: a high efficient way of PCR amplification of random DNA libraries in aptamer selection. *PLoS One* 6: e24910

Shi H, Colavin A, Lee TK, Huang KC (2017) Strain library imaging protocol for high-throughput, automated single-cell microscopy of large bacterial collections arrayed on multiwell plates. *Nat Protoc* 12: 429 – 438

Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S (2015) Cell-size control and homeostasis in bacteria. *Curr Biol* 25: 385 – 391

Uphoff S, Lord ND, Okumus B, Potvin-Trottier L, Sherratt DJ, Paulsson J (2016) Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. *Science* 351: 1094 – 1097

Vvedenskaya IO, Zhang Y, Goldman SR, Valenti A, Visone V, Taylor DM, Ebright RH, Nickels BE (2015) Massively systematic transcript end readout,

"MASTER": transcription start site selection, transcriptional slippage, and transcript yields. *Mol Cell* 60: 953 – 965

Wallden M, Fange D, Lundius EG, Baltekin Ö, Elf J (2016) The synchronization of replication and division cycles in individual *E. coli* cells. *Cell* 166: 729 – 739

Wang HH, Isaacs FJ, Carr PA, Sun ZZ, George X, Forest CR, Church GM (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460: 894 – 898

Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S (2010) Robust growth of *Escherichia coli*. *Curr Biol* 20: 1099 – 1103

Williams R, Peisajovich SG, Miller OJ, Magdassi S, Tawfik DS, Griffiths AD (2006) Amplification of complex gene libraries by emulsion PCR. *Nat Methods* 3: 545 – 550