Electronic Supplementary Material (ESI) for Integrative Biology.
This journal is © The Royal Society of Chemistry 2014

Integrative Biology

Supplemental Information for

Microfluidic cell sorter-aided directed evolution of a protein-based calcium ion indicator with an inverted fluorescent response

Yongxin Zhao[1], Ahmed S. Abdelfattah[1], Yufeng Zhao[1], Araya Ruangkittisakul[2], Klaus Ballanyi[2], Robert E. Campbell[1]*, D. Jed Harrison[1]*

[1]Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

[2]Department of Physiology, University of Alberta, Edmonton, Alberta T6G 2S2, Canada

*Correspondence: Correspondence related to microfluidics should be addressed to D.J.H. (jed.harrison@ualberta.ca) and correspondence related to new FP variants should be addressed to R.E.C. (robert.e.campbell@ualberta.ca).

Inventory of supplemental information

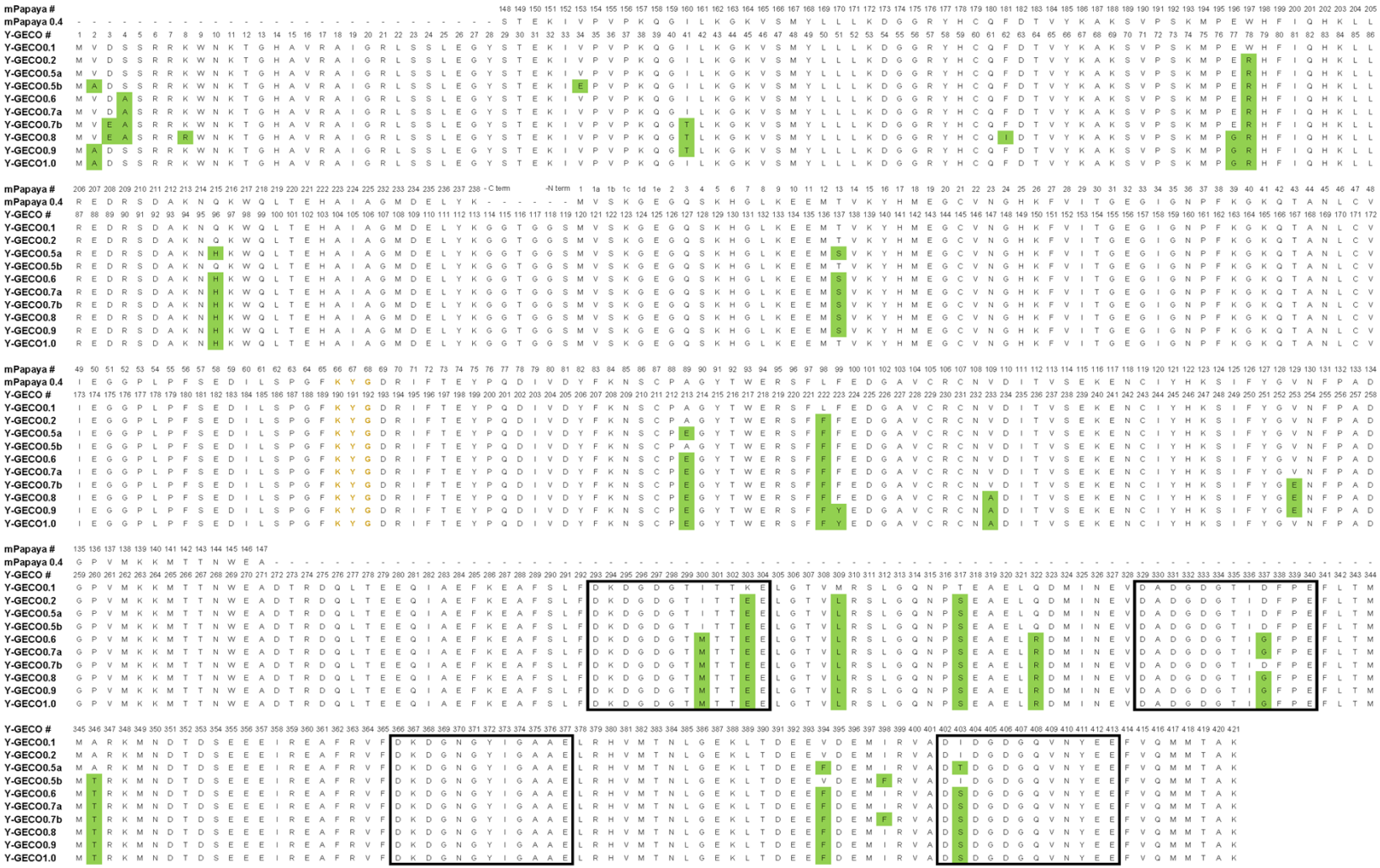
mPapaya #
mPapaya 0.4
Y-GECO #
Y-GECO0.1
Y-GECO0.2
Y-GECO0.5a
Y-GECO0.5b
Y-GECO0.6
Y-GECO0.7a
Y-GECO0.7b
Y-GECO0.8
Y-GECO0.9
Y-GECO1.0

Figure S1. Sequence alignment of Y-GECO1 and select Y-GECO variants. The KYG chromophore is highlighted in orange text. Positions of EF hands in CaM are boxed.

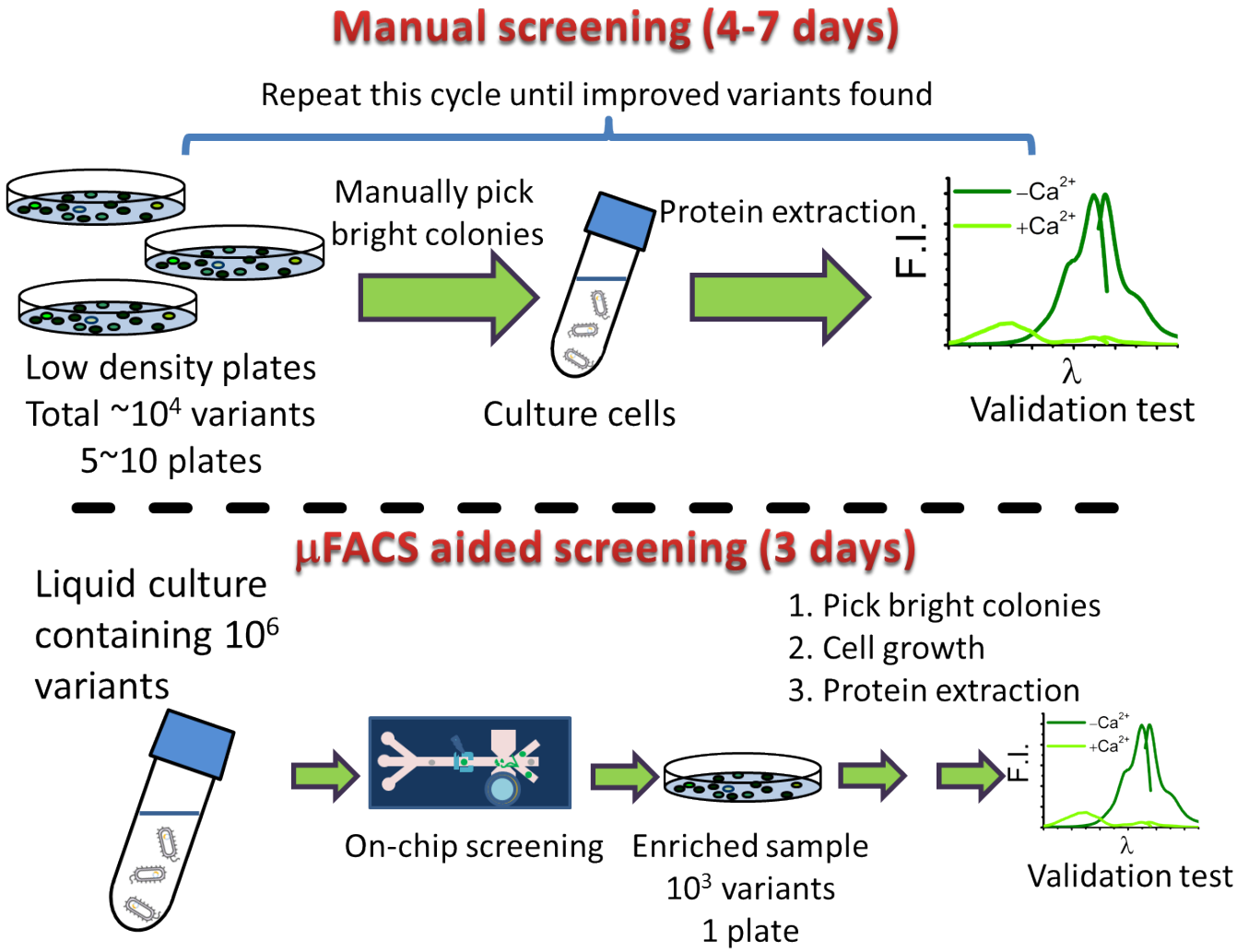

Figure S2. Schematic representation of screening procedures.

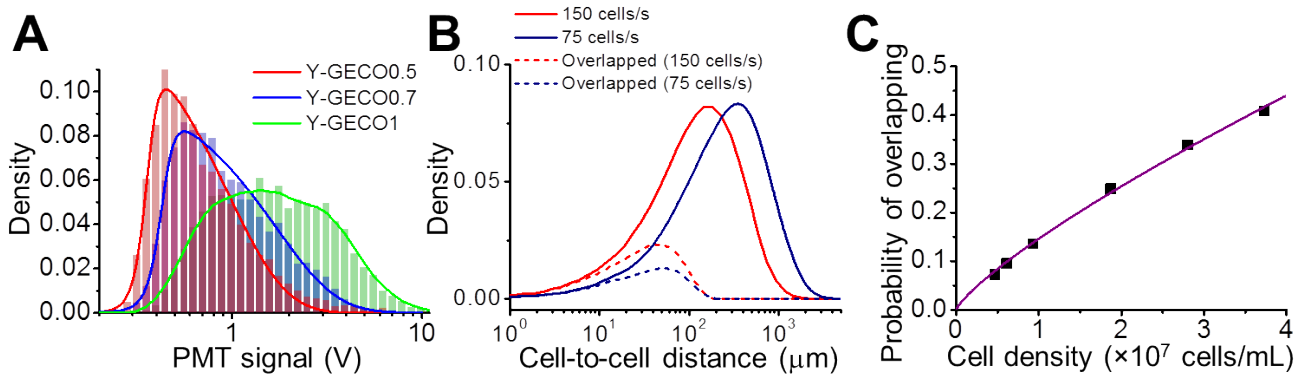

Figure S3. Supplemental analysis of μFACS. (A) Normalized histogram of distributions of PMT signals of three different clones, as detected by μFACS. (B) Simulated probability density distribution of cell-to-cell distance of an artificial sample in the main channel of μFACS with total flow rate 10.5 μL/min. Solid lines represent all the cells. Dash lines represent the co-sorted cells. (C) Fraction of overlapped cells as a function of relative screen rate to maximum throughput of μFACS. With total flow rate 10.5 μL/min, the maximum throughput of our μFACS is 150 cells/s.

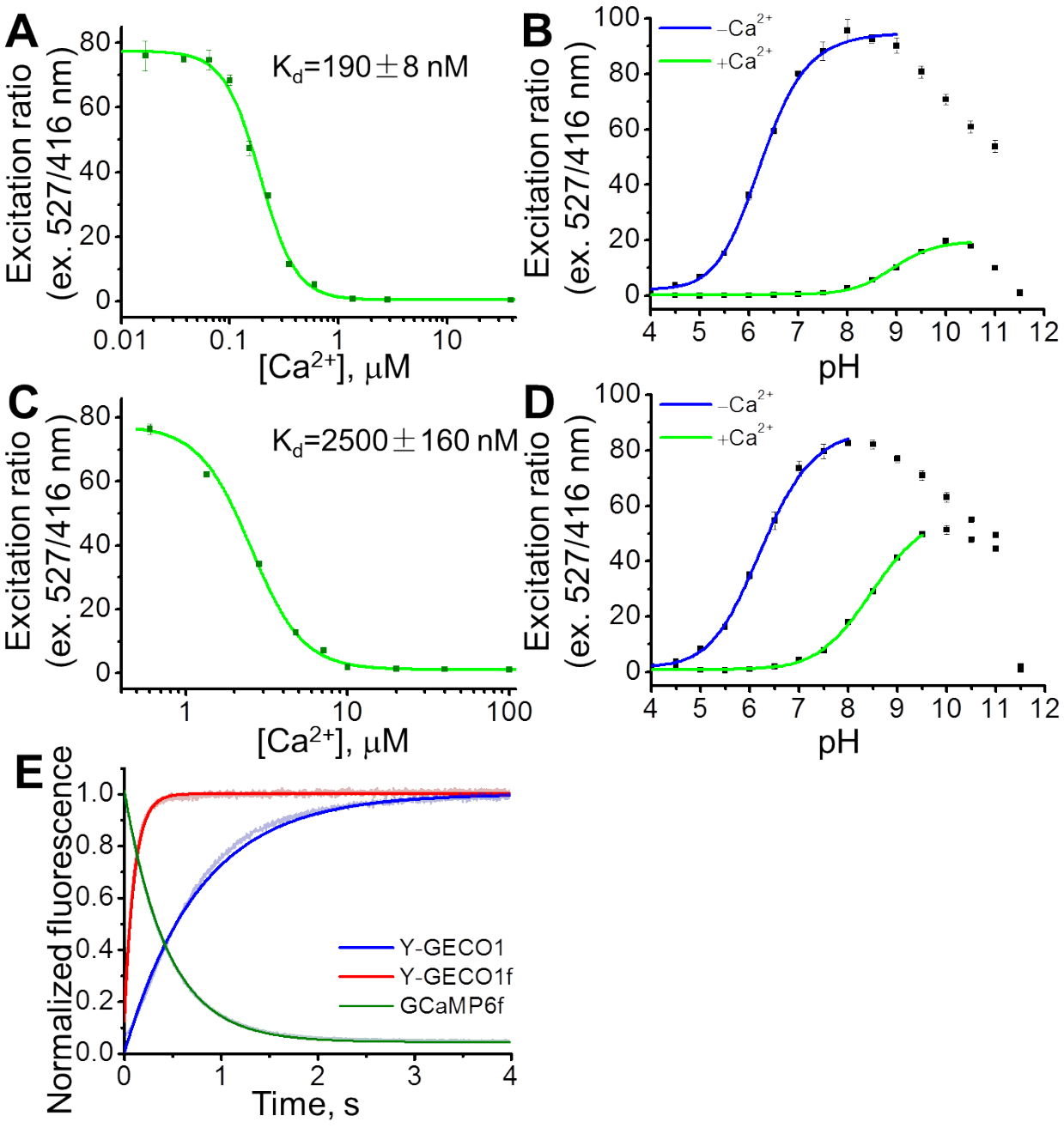

Figure S4. Characterization of *K_d*, *pKa* and kinetic of Y-GECO variants. (A,C) Ca^{2+} titration of (A) Y-GECO1 and (C) Y-GECO1f. (B,D) pH titration of (B) Y-GECO1 and (D) Y-

GECO1f with and without binding to Ca^{2+}. (E) Comparison of the off kinetic of Y-GECO1, Y-GECO1f and GCaMP6f.

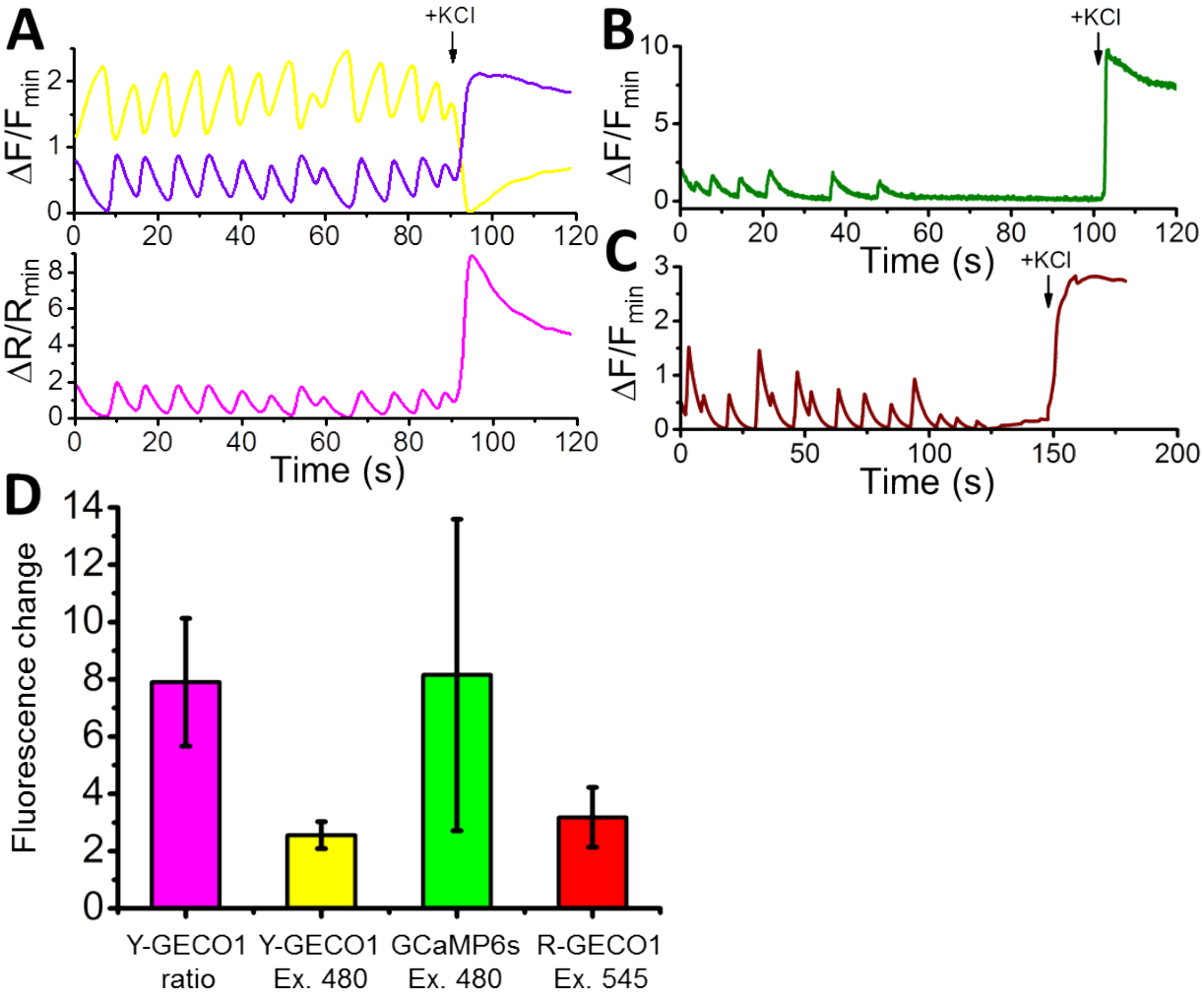

Figure S5. Performance of Y-GECO1 and select Ca^{2+} indicators in dissociated hippocampus neurons. (A) Imaging of spontaneous $[Ca^{2+}]_i$ oscillation in neurons expressing Y-GECO1. The sampling frequency is 0.8 Hz. Exposure: 70 ms with 12.5 % neutral density filter applied. (A, upper) Fluorescent cell traces acquired under 480 nm (yellow line) and 440 nm (violet line) excitations. (A, lower) Excitation ratios, 440 to 480 nm. (B) Imaging of spontaneous $[Ca^{2+}]_i$ oscillation in neurons expressing GCaMP6s, recorded at excitation 480 nm. The sampling frequency is 20 Hz. Exposure: 50 ms with 25% neutral filter applied. (C)

Imaging of spontaneous $[Ca^{2+}]_i$ oscillation in neurons expressing R-GECO1, recorded at excitation 545 nm and emission 590-650 nm. The sampling frequency is 25 Hz. Exposure: 40 ms without any neutral filter applied. (D) Comparison of fluorescence response of Y-GECO1, GCaMP6s and R-GECO1 in neurons upon KCl-induced depolarization in the presence of 50 mM KCl.

Table S1. Measure of sorting efficiency.

	Side outlet 1 (%)	Side outlet 2 (%)	Total (%)
No delay	44.0 (534/1213)*	0.0 (1/1317)	44.1
30ms delay	2.1 (47/2238)	0.0 (0/1183)	2.1
No actuation	0 (0/584)	0 (0/637)	0

*(number of cells detected in a side outlet / number of cells detected in main channel)

Table S2. Properties of Y-GECO1 and other variants.

Protein	Ca^{2+}	λ_{abs} (nm) with ε ($mM^{-1}\cdot cm^{-1}$) in parenthesis	λ_{em} with Φ in parenthesis[1]	Brightness[2] ($mM^{-1}\cdot cm^{-1}$)	pK_a[3]	Dynamic range	K_d' for Ca^{2+} (nM)
GCaMP6f	-	496 (4)	513 (0.10)	0.4	8.7 ± 0.3	50x	405 ± 2 (2.3 ± 0.2)
	+	496 (58)	513 (0.59)	34	6.3 ± 0.1		
Y-GECO0.8	-	410 (25); 523 (31)	513 (0.003, 0.55)	0.08, 17	6.4 ± 0.1	67x	180 ± 10 (2.4 ± 0.2)
	+	411 (38); 522 (9)	513 (0.03, 0.39)	1.1, 3.5	8.1 ± 0.1		
Y-GECO1 ΔG13 (Y-GECO1f)	-	410 (21); 523 (40)	538 (ND[4])	ND	6.3 ± 0.1	80x	2500 ± 160 (2.7 ± 0.2)
	+	413 (41); 522 (10)	538 (ND)	ND	8.5 ± 0.2		
Y-GECO1	-	410 (21); 523 (39)	538 (0.003, 0.76)	0.06, 30	6.3 ± 0.1	200x	190 ± 8 (2.6 ± 0.2)
	+	413 (39); 522 (6)	538 (0.05, 0.18)	1.1, 2.0	8.9 ± 0.1		

Note: The excitation ratio for Y-GECOs is defined as (fluorescence intensity for excitation at 524 nm and emission at 560 nm)/(fluorescence intensity for excitation at 413 nm and emission at 560 nm).[1]Quantum yield of

Y-GECO in the Ca^{2+}-bound state was measured for excitation at both 413 nm and 524 nm. [2]Brightness is defined as the product of ε and Φ. The brightness of Y-GECO in the Ca^{2+}-bound state for excitation at both 413 nm and 522 nm is provided. [3]The pK_a of Y-GECO is defined as the pH at which the dynamic range is 50% of maximum. [4]Not determined.

Table S3. Characterization of the Ca^{2+}-dependent fluorescence of Ca^{2+} indicators in HeLa cells. Cells were treated first with histamine (abb. His), then with Ca^{2+}/ionomycin (abb. Ca^{2+}), and finally with EGTA/ionomycin (abb. EGTA).

Protein (Excitation, nm)	n[1]	Maximum Ca^{2+} to minimum EGTA ratio	Maximum His to minimum His ratio	Maximum His to maximum Ca ratio
GCaMP6s (480)	39	20±7	11±3	0.9±0.3
Y-GECO1 (480)	57	6±1[2]	3.6±0.7[2]	0.6±0.1
Y-GECO1 (ratio[3])	57	13±3	9±2	0.6±0.3

[1]Number of individual transfected cells on which systematic calibration experiments were performed.

[2]The fluorescence change of Y-GECO1 is inverted when excited at 480 nm. The value is the reciprocal of the defined ratio in order to facilitate comparison.

[3]Excitation ratio: 440 nm to 480 nm.

Table S4. Oligonucleotides used in this work

Name	Sequence
RV_MluI_X_147mPapaya	GGCATGACGCGTRNVAGCTTCCCAATTAGTGGTCATTTTCTTCATTAC
FW_XhoI_X_148mPapaya	GCGATCCTCGAGBNYAGTACCGAGAAAATTGTGCCTGTTCCAAAGC
FW_XhoI_XX_148mPapaya	GCGATCCTCGAGBNYBNYAGTACCGAGAAAATTGTGCCTGTTCCAAAGC
FW_GGTGGS_mPapaya	GGAGGTACAGGCGGGAGTATGGTGAGCAAGGGCGAGGG
RV_GGTGGS_mPapaya	ACTCCCGCCTGTACCTCCCTTGTACAGCTCGTCCATGCCTGC
FW_XbaI_6His	GCGATGTCTAGAGGTTCTCATCATCATCATCATCATGGTATGGCTAGC
RV_CaM_stop_HindIII	GCGATGAAGCTTCTACTTCGCTGTCATCATTTGTACAAACTCTTCGTAGTTT
Fw_FCK_Gib_6His_YGECO	CTGGGGGCAGCGGGGGATCCACCATGGGTTCTCATCATCATCATCATCATGGTATGGC
RV_FCK_Gib_YGECO_stop	TCGATAAGCTTGATATCGAATTCTTACTTCGCTGTCATCATTTGTAC
Fw_del_13G	CGTAAGTGGAATAAGACACACGCAGTCAGAGCTATA
Fw_BamHI_Lyn	TGGGGATCCACCATGGGATGTATAAAGAGTAAAAGAAAAGATAACTTAA
Fw_Lyn_GECOs	ATGGGATGTATAAAGAGTAAAAGAAAAGATAACTTAAATGATGACGAAGATGATAAGGA TCTCGCCACAATG
Fw_BamHI_Kozak_6His	AAACAGGAGGAATTAAGCTTGGGATCCACCATGGGTTCTCATCATCATCATCATCATGG TATGGC
RV_CaM_stop_EcoRI	CGCGAATTCCTACTTCGCTGTCATCATTTGTAC

Movie S1. Wide-field imaging of Y-GECO1 in dissociated rat hippocampal neurons. Wide-field imaging of Y-GECO1 captured spontaneous $[Ca^{2+}]_i$ oscillation in two dissociated rat hippocampal neurons. The movie was made for three channels including yellow fluorescence under 440 nm excitation and 480 nm excitation as well as pseudocolored ratiometric change $\Delta R/R_{min}$ image (440 nm/480 nm). Each frame was acquired every 1 second. Scale bar: 30 μm. Movie frame rate: 10 frames/s.

Movie S2. Confocal imaging of Y-GECO1 in an organotypic brain slice. Neurons expressing human synapsin I promoter driven Y-GECO1 in an organotypic hippocampal slice was imaged using a confocal microscope. Prior to imaging, the slices were perfused with superfusate. The slice was then exchanged into superfusate containing 10 mM theophylline at the 10th second in the movie. Yellow fluorescence changes of Y-GECO reporting theophylline-evoked $[Ca^{2+}]_i$ oscillations were captured in the movie. Scale bar: 40 μm. Each frame was acquired every 0.5 s. Movie frame rate: 100 frames/s.

Supplemental experimental methods

Reagents

Synthetic DNA oligonucleotides used for cloning and library construction were purchased from Integrated DNA Technologies. The sequences of all oligonucleotides used in this work are provided in Table S3. *Pfu* polymerase (Fermentas) or AccuPrime™ *Pfx* SuperMix (Invitrogen) were used for non-mutagenic PCR amplifications in the buffer supplied by the respective manufacturer. *Taq* polymerase (New England Biolabs) in the presence of $MnCl_2$ (0.1 mM) was used for error-prone PCR amplifications. PCR products and products of restriction digests were routinely purified using preparative agarose gel electrophoresis followed by DNA isolation using the GeneJET gel extraction kit (Fermentas). Restriction endonucleases were purchased from Fermentas and used according to the manufacturer's recommended protocol. Ligations were performed using T4 ligase (Invitrogen). In some cases, Gibson Assembly (New England Biolabs) was used. Small-scale isolation of plasmid DNA was performed by GeneJET miniprep kit (Fermentas). The cDNA sequences for all Y-GECO variants and fusion constructs were confirmed by dye terminator cycle sequencing using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems).

Construction of Y-GECO gene libraries

The design of Y-GECO is based on well-established designs described by our group and others [1-5]. The initial cpmPapaya gene library was assembled by a two-part overlap extension PCR. The 5' piece used in the overlap extension was prepared by PCR amplification with a

mixture of two different forward primers (FW_XhoI_X_148mPapaya, FW_XhoI_XX_148mPapaya) and a single reverse primer (RV_GGTGGS_mPapaya). Primers FW_XhoI_X_148mPapaya and FW_XhoI_XX_148mPapaya contains an XhoI site and primer RV_GGTGGS_mPapaya encodes a cp linker (amino acid sequence GGTGGS) and an overlap region with primer FW_GGTGGS-mPapaya. The 3' piece for use in the overlap extension was prepared by PCR amplification with a single forward primer (FW_GGTGGS-mPapaya) and a different reverse primer (RV_MluI_X_147mPapaya). Primer RV_MluI_X_147mPapaya contains a MluI site. The PCR fragments were confirmed by agarose gel electrophoresis and purified. The full-length cpmPapaya gene library was assembled by overlap extension PCR using an equimolar mixture of primers FW_XhoI_X_148mPapaya, FW_XhoI_XX_148mPapaya and RV_MluI_X_147mPapaya together with a mixture of the 5' and 3' PCR fragments (1 μL each) as the template. The full-length product (approximately 700 bps) was purified by agarose gel electrophoresis and the doubly digested product was ligated into the XhoI and MluI sites of a modified pTorPE-G-GECO1.1 from which a second MluI site had been removed [1].

Error-prone PCR amplifications for construction of libraries of randomly mutated genes were performed using primers FW_XbaI_6His and RV_CaM_stop_HindIII. The PCR products were digested with XbaI and HindIII and ligated into the similarly digested pTorPE vector backbone. For site-directed mutagenesis or library construction by full or partial randomization of one or more codons, either the QuikChange Lightning Single or Multi kit (Agilent

Technologies) was used. Following ligation, electrocompetent *E. coli* strain DH10B cells was transformed with the library of gene variants and cultured overnight at 37 °C on 10-cm Petri dishes of LB-agar supplemented with 200 μg/mL ampicillin (Sigma) and 0.0020% (wt/vol) L-arabinose (Alfa Aesar).

Protein purification and *in vitro* spectroscopy

To purify Y-GECOs for *in vitro* spectroscopic characterization, the pTorPE plasmid harboring the variant of interest was first used to transform electrocompetent *E. coli* DH10B cells. Following selection on LB/ampicillin (200 μg/mL), single colonies were picked and used to inoculate 4 mL LB medium (200 μg/mL ampicillin only). Bacterial subcultures were shaken at 250 rpm and allowed to grow overnight at 37 °C. The next day, 1 mL of bacterial subculture was added into a modified TB rich medium (1L sterilized medium contains 20 g LB mix, 14 g tryptone, 7 g yeast extract, 9.2 g K_2HPO_4, 2.2 g KH_2PO_4 and 8 mL glycerol, pH was adjusted to 7.4). The cultures were shaken at 250 rpm and inoculated at room temperature for two days. Bacteria were harvested by centrifugation (10,000 g for 5 min), resuspended in 30 mM Tris-HCl buffer (pH 7.4), lysed by French press, and clarified by centrifugation at 13,000 g for 30 mins at 4 °C. Proteins were purified from the cell-free extract by Ni-NTA affinity chromatography (Agarose Bead Technologies). The buffer of purified proteins was exchanged into 10 mM MOPS, 100 mM KCl, pH 7.2. Absorption spectra were recorded on a DU-800 UV-visible spectrophotometer (Beckman) and fluorescence spectra were recorded on a Safire2 platereader. The response to Ca^{2+} is expressed as $(R_{max}-R_{min})/R_{min}$. For Y-GECO, R

= (*I* with 416 nm excitation)/(*I* with 527 nm excitation), where *I* = fluorescence intensity at 560 nm.

Standards for quantum yield determination were fluorescent protein mT-Sapphire[6] for the emission of Y-GECOs with 414 nm excitation and Citrine[7] for the emission of Y-GECOs with 524 nm excitation. The detailed procedure has been described previously [1]. The fluorescence emission spectra of each dilution of each standard and protein solution were recorded at excitation either 414 nm or 524 nm and the total fluorescence intensities obtained by integration. Integrated fluorescence intensity vs. absorbance was plotted for each protein and each standard. Quantum yields were determined from the slopes (*S*) of each line using the equation: $\Phi_{protein} = \Phi_{standard} \times (S_{protein}/S_{standard})$.

Extinction coefficients were determined by first measuring the absorption spectrum of each GECO in Ca^{2+}-free buffer (30 mM MOPs, 100 mM KCl and 10 mM EGTA at pH 7.20) and Ca^{2+}-buffer (30 mM MOPS, 100 mM KCl and 10 mM Ca-EGTA at pH 7.20). The concentrations of GCaMP6f were determined by measuring the absorbance following alkaline denaturation and assuming = 44,000 $M^{-1}cm^{-1}$ at 446 nm. For Y-GECOs, the protein concentration was determined by comparing the absorption peak for denatured Y-GECO to that of denatured mPapaya following alkaline denaturation of both proteins. Extinction coefficients of each protein were calculated by dividing the peak absorbance maximum by the concentration of protein.

pH titrations were carried out by diluting concentrated purified proteins into pH buffers containing 50 mM citrate, 50 mM Tris, 50 mM glycine, 100 mM NaCl, and either 5 mM $CaCl_2$ or 5 mM EGTA that were pre-adjusted to 14 different pH values between 4.50 and 11.50. pK_a values were determined from the inflection point of a sigmoid fit to fluorescence ratio versus pH.

Determination of Ca^{2+} K_d' was done according to the protocol described previously[1]. Briefly concentrated protein solution was diluted into a series of buffers which were prepared by mixing Ca^{2+}-saturated and Ca^{2+}-free buffers (30 mM MOPS, 100 mM KCl, 10 mM EGTA, pH 7.2, either with or without 10 mM Ca^{2+}) with free Ca^{2+} concentration ranges from 0 nM to 3,900 nM at 25 °C. The fluorescence ratio of Y-GECO in each solution was determined and plotted as a function of Ca^{2+} concentration. Experiments were performed in triplicate and the averaged data from the three independent measurements was fit to the Hill equation.

Ca^{2+}-dissociation kinetics was determined by stopped-flow photometry on an SX20 stopped-flow reaction analyzer (Applied Photophysics, Ltd., Leatherhead, UK). Excitation slit widths were set to 3 mm (equivalent to a wavelength bandwidth of 13.95 nm) and emission slit widths were set to 8 mm (equivalent to a wavelength bandwidth of 37 nm). Data were acquired using ProData SX software (Applied Photophysics, Ltd., Leatherhead, UK). The GECO indicator (in 30 mM MOPs, 0.2 mM $CaCl_2$ and 100 mM KCl) was rapidly mixed (1:1) with a buffered solution (30 mM MOPs, 10 mM EGTA and 100 mM KCl). The change in the

fluorescence signal during rapid mixing provided the dissociation k_{off}. All measurements were made at ambient temperature 20 °C.

Construction of mammalian expression plasmids

In this work, three types of mammalian expression plasmids were used. All genes were fully sequence verified before transfection. To construct the plasmids for validation and testing of Ca^{2+} indicators in HeLa cells, we cloned the our Y-GECO variants into a modified pcDNA vector that has been described previously[1]. The gene in the pBAD vector was first PCR amplified using primers Fw_BamHI_Kozak_6His and RV_CaM_stop_EcoRI. PCR products were gel purified and digested with BamHI and EcoRI. Purified DNA fragments were ligated into the modified pcDNA3 plasmid digested with BamHI and EcoRI and similarly purified. The ligation reaction was used for transformation of electrocompetent *E. coli* DH10B cells. Cells were plated on LB/agar supplemented with ampicillin and individual colonies were picked into 4 mL of LB/ampicillin following overnight incubation at 37 °C. Liquid cultures were shaken at 250 rpm and 37 °C for 12-15 h and then a small scale isolation of plasmid DNA was performed.

For expression in neurons, we cloned our Y-GECO variants into the BamHI/EcoRI sites of a lentivirus vector pJMK004 (FCK-ArchD95N-GFP, Addgene: 34616) or the BamHI/HindIII sites of an adeno-associated virus vector pAAV2-hSyn-R-GECO1. All the viral vectors were used directly for transfection without performing virus production. The FCK-ArchD95N-GFP vector contains a CaMKII promoter and a Woodchuck Hepatitis Virus Posttranscriptional

Regulatory Element (WPRE) after the 3' end of the open reading frame, and the pAAV2-hSyn-R-GECO1 vector contains a human synapsin I promoter and a WPRE after the 3' end of the open reading frame. For construction of FCK-Y-GECO plasmids, the Y-GECO DNA for ligation was first generated by PCR using forward primer Fw_FCK_Gib_6His_YGECO and reverse primer RV_FCK_Gib_YGECO_stop, purified as described before. The ligation was done by a Gibson Assembly reaction with the lentivirus vector digested by BamHI and EcoRI following the manufacturer's instruction. For construction of pAAV2-hSyn-Y-GECO plasmids, the Y-GECO DNA fragment was generated by PCR using forward primer Fw_FCK_Gib_6His_YGECO and reverse primer RV_CaM_stop_HindIII. In case of preparing pAAV2-hSyn-Lyn-Y-GECO plasmids, Forward primers Fw_BamHI_Lyn and Fw_Lyn_GECOs and reverse primer RV_CaM_stop_HindIII were used to generate DNA fragment via two steps of PCR. The DNA fragments were then purified, digested and ligated into the cut adeno-associated virus vector digested by BamHI and HindIII.

HeLa cell culture

HeLa cells (40-60% confluent) on home-made 35 mm glass bottom dishes were transfected with 1 μg of plasmid DNA and 2 μL Turbofect (Thermo Scientific) or 3 μL lipofectamine 2000 (Invitrogen) according to the instruction of respective manufacturers. After 3 h incubation the media was exchanged to DMEM with 10% fetal bovine serum (FBS) and the cells were incubated for an additional 24 h at 37 °C in a CO_2 incubator. Immediately prior

to imaging, cells were washed twice with Hanks balanced salt solution (HBSS) and then 1 mL of 20 mM HEPES buffered HBSS was added.

Preparation of rat hippocampal neuron culture

Dissociated hippocampal cells from Sprague-Dwaley rats of embryonic day 18 (E18) stored in Hibernate® EB Complete Media were purchased from BrainBits LLC. The cells were grown on a homemade 35-mm glass bottom dish containing NbActiv4 (BrainBits LLC) supplemented with 2% FBS, penicillin-G potassium salt (50 units/mL), and streptomycin sulfate (50 μg/mL). On the 4th day *in vitro* (DIV-4), half of the culture medium was replaced with fresh media. Neuronal cells were transfected at DIV-7 with chosen plasmids using Lipofectamine 2000 or GenJet (SignaGen Laboratories) according to the manufacturer's instruction. Cells were imaged 2-4 days after transfection.

Supplemental text

Measurement of the distribution of expression levels for a single clone by μFACS

To measure the distribution of expression levels for a single clone, cells from a single fluorescent colony were collected and suspended in M9 buffer as described in experimental procedures. The cell suspension (around 5×10^6 cells/mL) was then loaded in a μFACS with a total flow rate 10.5 μL/min. The PMT signals were recorded continuously for 10 seconds and the recording was repeated 40 times. All 40 10-s spectra were processed by a customized Matlab program to extract the fluorescent peaks for statistical analysis. Three different Y-GECO constructs obtained from different stages of directed evolution were tested separately. The normalized distributions of cell fluorescence of three clones are plotted in Figure S3A. As shown in Figure S3A, the expression level of a single clone is fairly broad, but the upper range of fluorescence brightness reflects the absolute brightness of the variant.

Measurement of sorting efficiency using piezoelectric actuation

To measure the sorting efficiency, a second illumination and detection point was employed at one of the side collection channels. The signals of sorted cells in the side collection channel were recorded by a second PMT detector. The sorting efficiency was determined by the following equation:

$$Sorting\ efficiency = \frac{number\ of\ signals\ detected\ by\ PMT\ \#2}{number\ of\ signals\ detected\ by\ PMT\ \#1} \times 100\%$$

We measured the sorting efficiency under typical parameters used for screening Y-GECO. The results are summarized in Table S1. The total flow rate was 10.5 μL/min for all the tests. Cells were sorted into one side collection channel, and the sorting efficiency observed was around 44%. Without triggering the piezoelectric actuator, there is essentially no cell flow into the side collection channel, limiting contamination problems. We also measured the sorting efficiency by applying a 30 ms long delay that is 14 times longer than the system response time (2.2 ms). As expected, the sorting efficiency drops to 2% with inaccurate timing for sorting.

Considerations for setting the throughput of μFACS

As discussed in the main text, the cell density of samples should be adjusted to ~2×10^7 cells/mL, so that there is only an average of one or less cell per excitation volume (1.1 nL). This is a simplified estimation and, statistically, there will be excitation volumes that do contain more than one cell, regardless of the cell density. If there is more than cell in the same excitation volume, we refer to them as “overlapped”. The portion of overlapped cells in the chip with a given cell density can be estimated by numerical analysis.

To estimate the portion of overlapped cells with a given cell density, we first need to establish a model. We consider cells are distributed randomly in a line with fixed length. The cell density determines how many cells are distributed along this line. We consider two cells might occupy the same excitation volume if the distance is less than 150 μm. For these two

overlapped cells, the distance determines the probability of co-sorting $p_{co-sorting}$, as indicated by the following equation:

$$p_{co-sorting} = \frac{150 - d}{150 + d} \, (d \leq 150 \, \mu m) \tag{1}$$

Where *d* is the cell-to-cell distance (μm). With this model, we can perform simple numerical analysis and statistically analyze the distribution of cell-to-cell distance (Figure S3B and C).

As shown in Figure S3B, the distribution of cell-to-cell distance in the chip could be satisfactorily fit to a Pearson type IV distribution. With a cell density of 1.87×10^7 cells/mL, the probability of co-sorting for each piezoelectric actuation is around 25% and is reduced to 13% with 50% the cell density. A 13% to 25% probability of co-sorting is acceptable in our study, as we have a secondary screen for further isolation of brightest variants. Figure S3C summarizes the probability of co-sorting as a function of cell density of the sample. Figure S3B and S3C suggest that co-sorting is one of the main factors that affect the enrichment efficiency and it could be minimized by reducing the cell density of the sample. With reduced cell density of the sample, the throughput can be maintained by increasing the total flow rate at the cost of lower SNR for detection.

Supplemental references

1 Y. Zhao, S. Araki, J. Wu, T. Teramoto, Y.-F. Chang, M. Nakano, A. S. Abdelfattah, M. Fujiwara, T. Ishihara, T. Nagai, R. E. Campbell, An Expanded Palette of Genetically Encoded Ca^{2+} Indicators. *Science* 2011, *333*. 1888-1891.

2 H. J. Carlson, R. E. Campbell, Circular permutated red fluorescent proteins and calcium ion indicators based on mCherry. *Protein Eng. Des. Sel.* 2013, *26*. 763-772.

3 J. Akerboom, N. Carreras Calderón, L. Tian, S. Wabnig, M. Prigge, J. Tolö, A. Gordus, M. B. Orger, K. E. Severi, J. J. Macklin, R. Patel, S. R. Pulver, T. J. Wardill, E. Fischer, C. Schüler, T.-W. Chen, K. S. Sarkisyan, J. S. Marvin, C. I. Bargmann, D. S. Kim, S. Kügler, L. Lagnado, P. Hegemann, A. Gottschalk, E. R. Schreiter, L. L. Looger, Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics. *Front Mol. Neurosci.* 2013, *6*.

4 T. Nagai, A. Sawano, E. S. Park, A. Miyawaki, Circularly permuted green fluorescent proteins engineered to sense Ca^{2+}. *Proc. Natl. Acad. Sci. U.S.A.* 2001, *98*. 3197-3202.

5 J. Nakai, M. Ohkura, K. Imoto, A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein. *Nat. Biotechnol.* 2001, *19*. 137-141.

6 H.-w. Ai, K. L. Hazelwood, M. W. Davidson, R. E. Campbell, Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat. Methods* 2008, *5*. 401-403.

7 O. Griesbeck, G. S. Baird, R. E. Campbell, D. A. Zacharias, R. Y. Tsien, Reducing the Environmental Sensitivity of Yellow Fluorescent Protein: MECHANISM AND APPLICATIONS. *J. Biol. Chem.* 2001, *276*. 29188-29194.