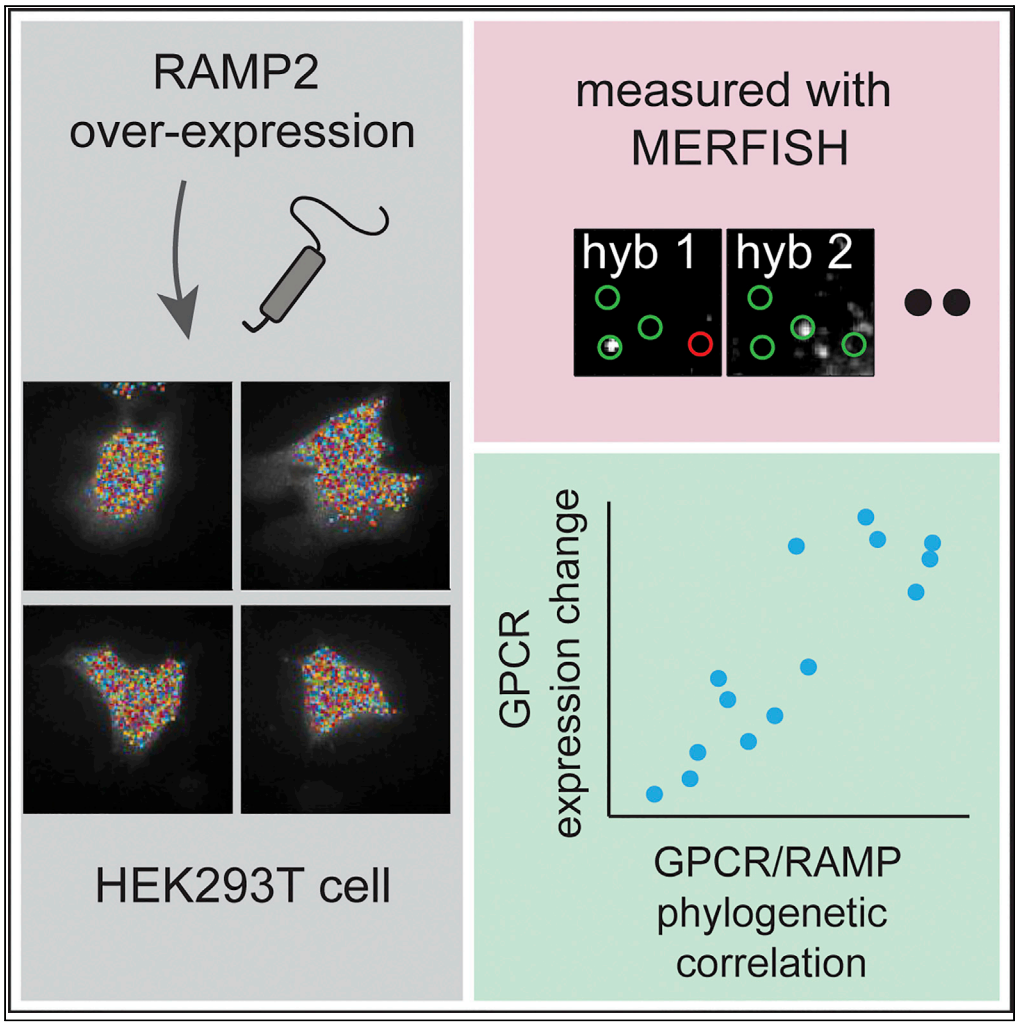
iScience

CellPress

Article

Detection of Concordance between Transcriptional Levels of GPCRs and Receptor-Activity-Modifying Proteins



Shahar Barbash,
Torbjörn Persson,
Emily Lorenzen,
Manija A. Kazmi,
Thomas Huber,
Thomas P. Sakmar

hubert@rockefeller.edu (T.H.)
sakmar@rockefeller.edu (T.P.S.)

HIGHLIGHTS

A multiplexed fluorescence *in situ* hybridization (MERFISH) approach was used

GPCR and RAMP2 transcriptional levels were measured in single cells

Transcriptional levels were correlated as predicted by phylogenetic analysis

Results suggest that RAMPs and GPCRs globally interact with functional consequences

DATA AND SOFTWARE AVAILABILITY

GSE122633

Barbash et al., iScience *11*, 366–374
January 25, 2019
https://doi.org/10.1016/j.isci.2018.12.024

iScience

CellPress

Article

Detection of Concordance between Transcriptional Levels of GPCRs and Receptor-Activity-Modifying Proteins

Shahar Barbash,[1] Torbjörn Persson,[1,2] Emily Lorenzen,[1] Manija A. Kazmi,[1] Thomas Huber,[1,*] and Thomas P. Sakmar[1,2,3,*]

SUMMARY

A recent phylogenetic analysis showed global co-evolution of G protein-coupled receptors (GPCRs) and receptor-activity-modifying proteins (RAMPs) suggesting global interactions between these two protein families. Experimental validation of these findings is challenging because in humans whereas there are only three genes encoding RAMPs, there are about 800 genes encoding GPCRs. Here, we report an experimental approach to evaluate GPCR-RAMP interactions. As a proof-of-concept experiment, we over-expressed RAMP2 in HEK293T cells and evaluated the effect on the transcriptional levels of 14 representative GPCRs that were selected based on the earlier phylogenetic analysis. We utilized a multiplexed error-correcting fluorescence *in situ* hybridization (MERFISH) method to detect message levels for individual GPCRs in single cells. The MERFISH results showed changes in GPCR message levels with RAMP2 over-expression in a concordant pattern that was predicted by the earlier phylogenetic analysis. These results provide additional evidence that GPCR-RAMP interactions are more widespread than previously appreciated and that these interactions have functional consequences.

INTRODUCTION

G protein-coupled receptors (GPCRs) participate in a wide range of basic molecular processes and are highly druggable therapeutic targets (Venkatakrishnan et al., 2013). Receptor-activity-modifying proteins (RAMPs) have been shown to interact with several GPCRs and affect their ligand specificity, cellular trafficking, and post-translational modification (McLatchie et al., 1998; Parameswaran and Spielman, 2006; Weston et al., 2015). For example, the calcitonin receptor-like receptor (CALCRL), arguably the most studied GPCR in this context, interacts with RAMP1 in a way that facilitates the receptor's transport to the plasma membrane (Parameswaran and Spielman, 2006) and determines its ligand specificity (Hay and Pioszak, 2016; Hay et al., 2006). Recently, we compared estimates of GPCR-RAMP co-evolution with those for established interacting protein pairs such as subunits of receptor complexes and endogenous protein ligands with their receptors. We found global co-evolution between GPCRs and RAMPs, suggesting that the interactions between the protein products of the two gene families are more widespread than previously presumed (Barbash et al., 2017).

Previous studies showed that GPCR activation modulates RNA stability (Tholanikunnel and Malbon, 1997) and general transcriptional programs (Lefkowitz and Shenoy, 2005) as well as specific transcriptional programs for the activated receptors (Sharp et al., 2013; Tsao et al., 2001). Based on these observations, we hypothesized that there might be some degree of correlation between RAMP and GPCR expression changes and that we could use such a correlation to validate our previous co-evolution analysis experimentally. Recent developments in highly multiplexed RNA profiling enable single-molecule measurements with error detection and correction in single cells (Chen et al., 2015). We modified this technique to enable detection of shorter genes when compared with the original MERFISH method (see Transparent Methods). We applied our modified MERFISH method to a subset of target genes in HEK293T cells over-expressing RAMP2 and proper controls for correlated detection of 14 GPCR messages. We compared the across-cells expression change data with the results from our previous global phylogenetic analysis. We conclude that the levels of GPCR-RAMP interactions determined by RNA expression and the interactions determined by co-evolution analyses correspond. The concordance of the phylogenetic and expression measurements gives further confidence in the "global GPCR-RAMP interaction" hypothesis. In summary, we establish a

[1]Laboratory of Chemical Biology and Signal Transduction, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

[2]Department of Neurobiology, Care Sciences and Society, Division for Neurogeriatrics, Center for Alzheimer Research, Karolinska Institutet, 141 57 Huddinge, Sweden

[3]Lead Contact

*Correspondence: hubert@rockefeller.edu (T.H.), sakmar@rockefeller.edu (T.P.S.)

https://doi.org/10.1016/j.isci.2018.12.024

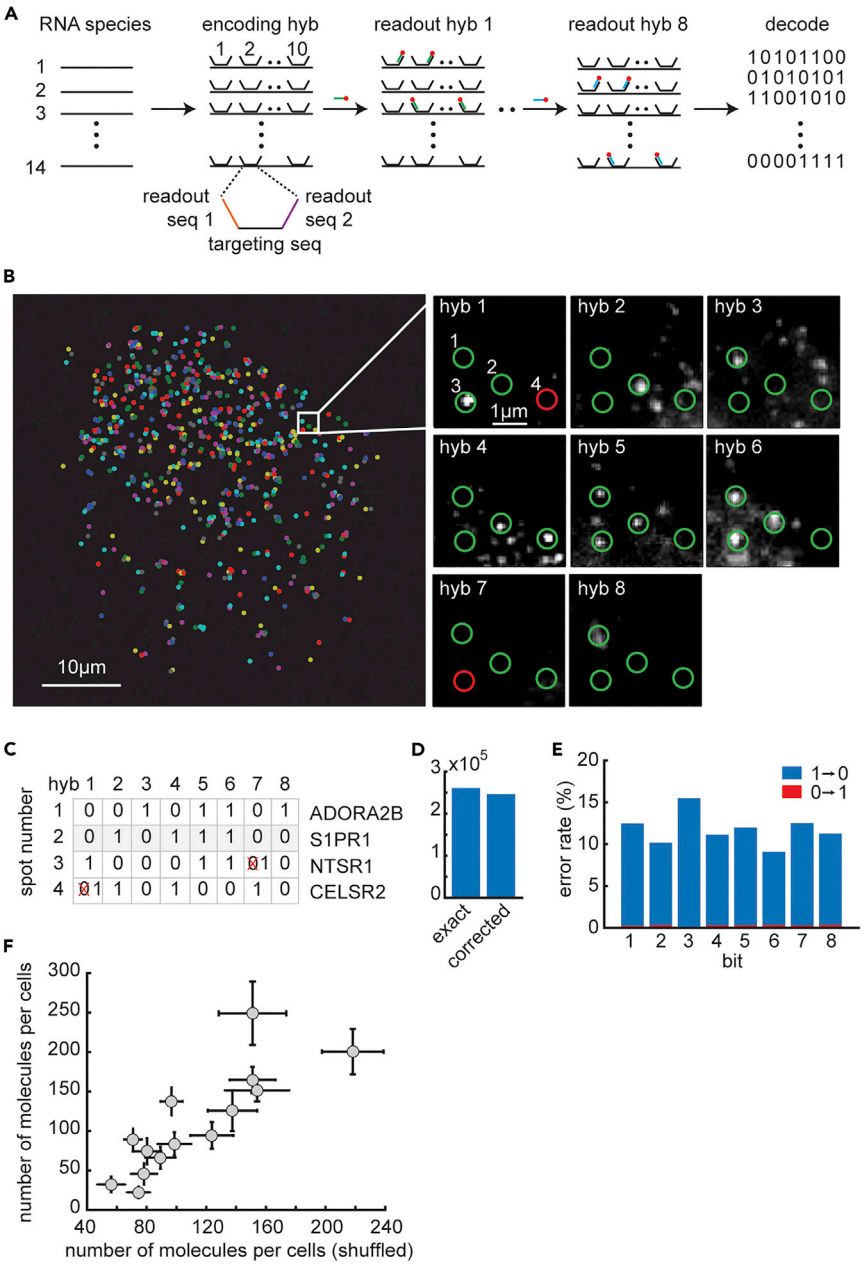


Figure 1. Design and Characterization of a MERFISH Tailored for GPCRs

(A) Hybridization scheme, as explained in Transparent Methods. Briefly, 10 encoding probes per gene were hybridized followed by a series of eight hybridization rounds in which fluorescent probes were flowed across the sample, hybridized with the corresponding encoding probes, imaged, and then bleached. Last, genes were decoded by their built-in code words.

(B) All detected single molecules in an example cell are shown colored based on gene identity. Four localizations are shown in the enlarged subregion (white square) in which single molecule was detected based on the code book, with error correction, across the eight hybridization rounds. Green circles indicate signals that correspond to the expected bit in the code word of a gene, whereas red circles indicate corrected bits. This correction was enabled due to Hamming distance 4 between the code words, as explained in Transparent Methods.

(C) Summary of the detection results shown in (B) for each location and the corresponding decoded genes.

(D) The total number of RNAs decoded without (exact) and with (corrected) error correction.

(E) The error rate for each bit. The error rate is calculated as the fraction of measured bar codes that contain a given bit flip. Both 1-to-0 error rates (blue) and 0-to-1 error rates (red) are shown for each bit.

Figure 1. *Continued*

(F) Scatterplot for the number of molecules detected per cell for each of the 14 GPCRs in the original "code word" assignment (i.e., the order of hybridization rounds during imaging, see Transparent Methods) versus the shuffled version. Error bars are SEM across cells in each of the conditions. Pearson's correlation coefficient for the scatter plot is 0.83 and correlation p value <0.01.

potential experimental approach to validate earlier phylogenetic results and give further support for GPCR-RAMP interactions based on RNA expression patterns in HEK293T cells.

RESULTS

Designing and Characterizing MERFISH for GPCRs

The direct experimental testing of all the possible pairwise interactions among the three RAMPs and over 800 GPCRs is not currently feasible. However, we wanted to devise an experimental method to begin to validate the hypothesis that there exists a global interaction network between RAMPs and GPCRs. Therefore, we adopted the MERFISH method and applied it to test a selected tractable subset of GPCRs expressed in HEK293T cells. HEK293T cells are widely used for the pharmacological characterization of GPCR signaling pathways. Approximately 61 individual GPCRs are known to be expressed in HEK293 cells in culture (Atwood et al., 2011). We used several criteria to select 14 representative GPCRs (Table S1) for examination in the MERFISH assay: (1) homogeneous distribution across the phylogenetic measures (Barbash et al., 2017) to represent either high or low phylogenetic association with RAMP2, (2) known expression in HEK293T cells, and (3) satisfactory message length for hybridization of enough detecting probes (10 for each gene) to establish a sufficient signal-to-noise ratio.

We adopted and modified the MERFISH technique described earlier (Chen et al., 2015) to enable detection of shorter mRNA transcripts typical for GPCRs. The original MERFISH method used 96 probes per transcript, limiting the method in practice to genes of at least 3,000-nucleotide (nt) in length. As family A (rhodopsin-like) GPCRs are typically approximately 350 amino acids in length, their transcripts are about 1,050 nt long and are too short for standard MERFISH. Briefly, a set of 30-nt-long orthogonal DNA probes was designed with standard probe optimization methods (see "Probe Design and Synthesis" in Transparent Methods for criteria) for the chosen 14 GPCRs with 10 probes per GPCR gene. These probes have 30-nt-long unhybridizing flanking tails on both ends of the target region that hybridize with 30-nt-long fluorescent probes (see Transparent Methods, Figure 1A for hybridization scheme and Tables S1 and S2 and Data S1 for full probe sequences). The modification described also allowed for more efficient detection of a relatively low number of genes.

The fluorescent probes in cultured cells were detected with a Nikon Ti-E inverted microscope in a series of hybridization and bleaching rounds that build up a series of detection and no-detection readouts for each location in the observed field. This series is later treated as a binary word of eight bits where 1 and 0 denote detection and no detection and each bit corresponds to a single hybridization round. See Figure S1A for an example of raw image data and Figure S1B for analysis of bleaching efficiency. Hybridization images for each cell field were registered based on signals from carboxylate-modified microspheres (see "Sample Preparation" in Transparent Methods and Figure S1C). Series signals were detected by identifying objects of higher intensity compared with their surroundings, followed by size filtration, clustering of identical neighboring objects, and a second round of size filtration (see "Image Analysis" in Transparent Methods). Next, these series were matched against the predicted pattern for each gene, while allowing for error detection and correction based on the probe library design (see "Probe Design and Synthesis" in Transparent Methods and Figures 1B and 1C). About one-half of RNA molecules were decoded without error correction, and about one-half were decoded with error correction (Figure 1D). The measured error rate (the event of a bit flip) across the hybridization rounds was 10%–15% (Figure 1E). Example cells overlaid with the 14 detected GPCRs are shown in Figure S2A. We performed a permutation version of the order of hybridizations ("shuffled" version) and measured similar expression levels, validating the library design, imaging, and analysis steps (Figure 1F). We also measured the averaged signal intensity due to sample autofluorescence and probe non-specific binding and found the value to be less than 5% of the averaged signal intensity in any of the hybridization rounds, which we considered satisfactory.

Over-expression of RAMP2 in HEK293T Cells

In our previous work (Barbash et al., 2017) we showed that RAMP1 and RAMP3 co-evolved with similar GPCRs and have structural resemblance at the extracellular region, which is required for transport of

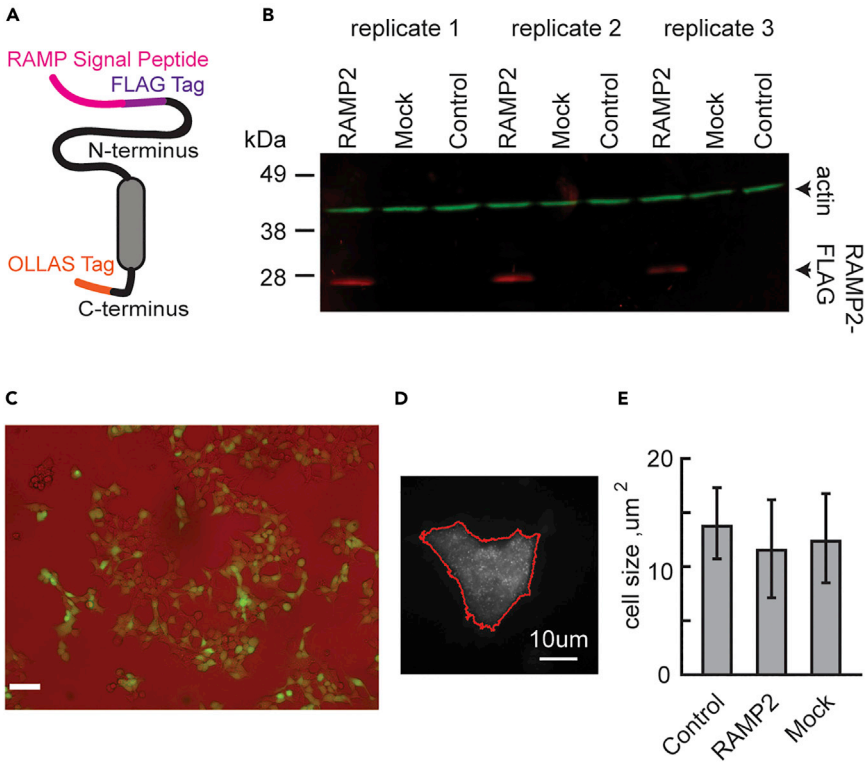


Figure 2. Over-expression of RAMP2 in HEK293T Cells

(A) HEK293T cells were transfected with a RAMP2 DNA construct that encoded a FLAG epitope tag on the N-terminal end and an OLLAS epitope tag on the C-terminal end. The entire native sequence of RAMP2 except for the initiator Met and predicted cleavable signal peptide was included between the epitope tags. The DNA construct was in a pcDNA3.1+ vector.

(B) Immunoblot of transfected HEK293T cells using antibodies against FLAG (red) and actin (green). RAMP2, cells transfected with RAMP2 DNA construct; MOCK, cells transfected with empty pcDNA3.1 plasmid; control, non-transfected cells.

(C) HEK293T cells transfected with a pIRES2-EGFP vector. Bright field is represented by red channel, and GFP detection is represented by green channel; 48% of the cells were GFP-positive 48 h post transfection. Scale bar, 1 mm.

(D) Borders of cells were traced with the Moore-Neighbor tracing algorithm, based on general localized intensities and their area calculated.

(E) Bar graph of cell area (mean, SEM across all cells in condition) shows no significant difference in cell size across conditions. One-Way-ANOVA p value >0.05.

RAMP1-CALCRL to the plasma membrane. These observations suggest redundancy at the functional level. In other words, RAMP2, presumably, is distinct in its set of interacting GPCRs. Under over-expression of either RAMP1 or RAMP3, there would have likely been a compensation of one on the effect of the other. Therefore, it would have been difficult to expect a RAMP1- or RAMP3-specific effect on GPCRs' expression, and so difficult to predict that the pattern of GPCR expression change would correlate with the observed phylogenetic pattern. RAMP2 over-expression, on the other hand, does not suffer from this hurdle. For this reason, we designed and conducted RAMP2 over-expression experiments.

The steps described above were performed under three experimental conditions: non-transfected (control) HEK293T cells, cells transfected with a FLAG tag construct expressing RAMP2 (see Figure 2A for construct schematic), and cells transfected with an empty construct (mock) of the same vector backbone (pcDNA 3.1) as the one for the RAMP2. Three experimental replicates were done for each condition, and the number of cells analyzed in each one was 47 cells in control experiments, 37 cells for RAMP2 over-expression, and 53 cells for mock (see Transparent Methods for criteria for cells that were filtering out during imaging). RAMP2 expression across the three conditions is shown in Figure 2B. Transfection efficiency was estimated by HEK293T cells transfected with a pIRES2-EGFP vector (Figure 2C). No significant cell size difference was observed among the conditions tested (Figures 2D and 2E). We compared

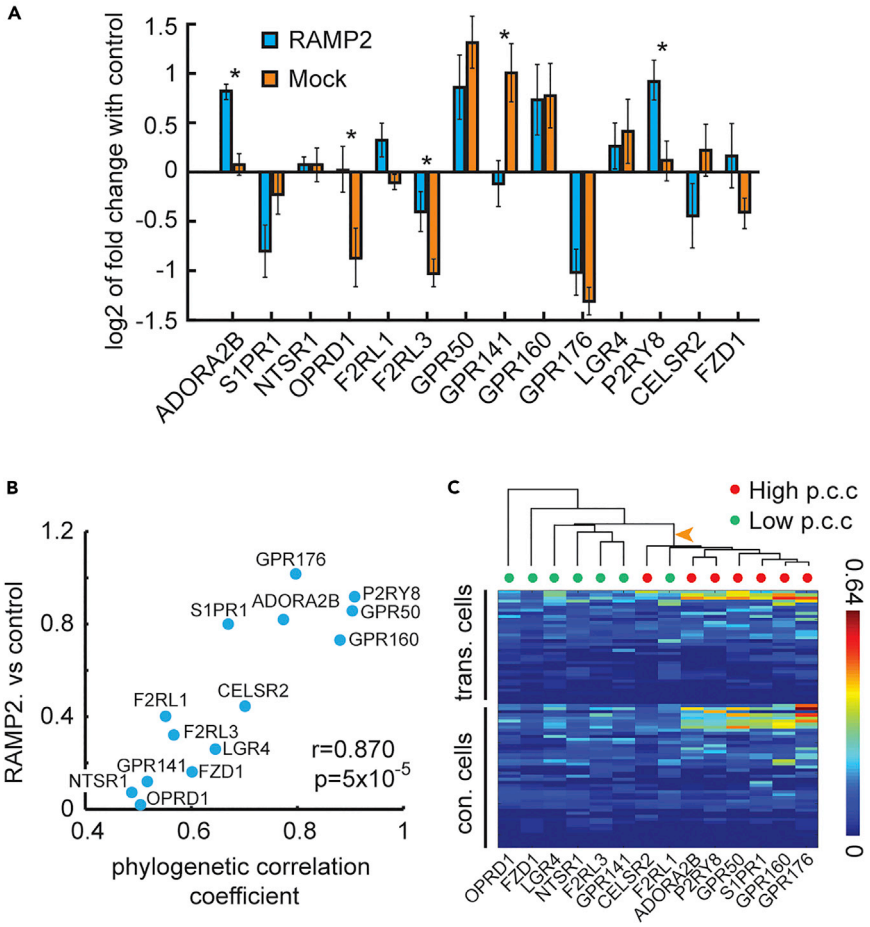


Figure 3. GPCR-RAMP2 Co-expression Interaction in HEK293T Cells

(A) Bar graph (mean, SEM) of the log2 transformation of fold change of GPCR expression between HEK293T cells transfected with RAMP2 and control cells (blue) and between Mock transfection and control cells (orange) for 14 GPCRs detected with the MERFISH method. Asterisks denote significant t test p values corrected for multiple comparisons with Bonferroni.

(B) Scatterplot of log2 absolute value of fold change of GPCR expression between HEK293T cells transfected with RAMP2 and control cells (y axis) versus phylogenetic correlation coefficient between RAMP2 and each of the GPCRs (x axis). Pearson's r = 0.87 and p value = $5*10^{-5}$, also shown in figure.

(C) Hierarchical clustering tree (dendrogram) based on GPCR expression distance (euclidean) between RAMP2 transfection cells and non-transfected control cells shows significant clustering of the top seven phylogenetically associated GPCRs (red circles) on the branch marked with an orange arrow (two-tailed Fisher's exact test for the clustering, p = 0.0047). Color bar represents the density of identified RNA molecules (number of molecules/pixel). p.c.c., phylogenetic correlation coefficient.

relative expression values of the 14 GPCRs with two available RNA-sequencing (RNA-seq) datasets (GSE53386 and GSE60559) of HEK293T and observed high correlations (Figure S2B).

GPCR-RAMP2 Co-expression Interaction in HEK293T Cells

These expression data were compared to the phylogenetic measures on a cell-by-cell basis and on a population level. On the population level, for each GPCR we calculated the absolute value of log2 for the fold change of expression between RAMP2 transfection and control, and similarly between mock transfection and control (Figure 3A). An absolute value of the log2 for the fold change was then plotted against the phylogenetic mammalian measures of the 14 GPCRs. Specifically, the absolute value was used because there was no reason to favor a prediction for either up- or down-regulation. We tested the most straightforward hypothesis that if a gene pair interacts, as estimated by phylogenetic analysis,

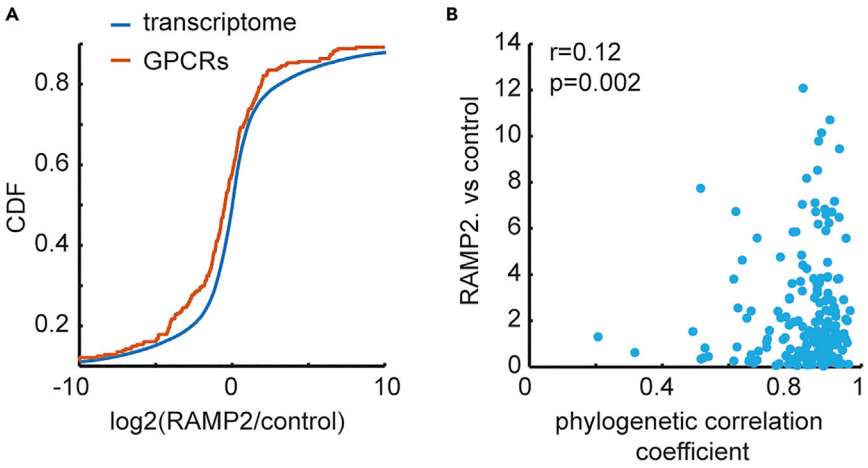


Figure 4. Global GPCR Expression Change upon RAMP2 Over-Expression

(A) Cumulative Distribution Function (CDF) of log2 of fold change between the average across three samples of RAMP2 over-expression and three control samples. CDF is shown separately for the whole transcriptome (red) and for all GPCR transcripts (red). Kolmogorov-Smirnov p value for difference between distributions: 0.048.

(B) Scatterplot of log2 absolute value of fold change of all GPCR expression, based on RNA-seq-derived Fragments Per Kilobase Million (FPKM) values, between HEK293T cells transfected with RAMP2 and control cells (y axis) versus phylogenetic correlation coefficient between RAMP2 and each of the GPCRs (x axis). Pearson's r = 0.12 and p value = 0.002, also shown in figure.

then the expression of one member is likely to affect the expression of the other in either the upward or downward direction.

A significant correlation between mammalian phylogenetic measures and MERFISH expression was observed for the RAMP2 transfection, but not for the mock transfection (Figures 3B and S3A–S3C). On the cell-by-cell level, we first split the GPCR group for two subgroups equal in size (seven each) based on their phylogenetic correlation with RAMP2, which is equivalent to setting the phylogenetic correlation coefficient threshold at 0.65. Next, a dendrogram based on between-cells expression distances was built, and the members of the two subgroups were marked on it (Figure 3C). The dendrogram shows that the top RAMP2 phylogenetic-associated GPCRs changed in a similar manner, and independently from, the other group upon RAMP2 over-expression, which suggests a common effector. This expression-based clusterization was neither observed for mock versus control nor for mock versus RAMP2 over-expression (Figure S3D). Taken together, these data show measurable interaction for highly phylogenetically associated gene pairs.

Among the GPCRs tested, five showed a significant change of expression in MERFISH upon RAMP2 over-expression (Figure 3A). These were the class A GPCRs: ADORA2B, OPRD1, F2RL3, GPR141, and P2RY8. Based on our previous phylogenetic analysis, all five GPCRs have relatively high phylogenetic correlation coefficients with all three RAMPs. In addition, F2RL3, GPR141, and P2RY8 also have relatively high coexpression correlation coefficients with RAMP2 and RAMP3. Last, ADORA2B, OPRD1, F2RL3, and P2RY8 are expressed at high levels across human tissues, whereas GPR141 is expressed at moderate levels (Expression Atlas; https://www.ebi.ac.uk/gxa/home/).

To examine GPCR expression change upon RAMP2 over-expression for more GPCRs, we performed whole-exome expression profiling by RNA-seq for the same three experimental conditions as before (see Methods for details). We found a significant global shift of GPCR expression toward down-regulation compared with the rest of the transcriptome (Figure 4B and S4). A comparable expression shift was not observed for mock transfection (Figure S5A). Next, we examined the concordance of global GPCR expression as measured by RNA sequencing with the phylogenetic correlation coefficient. We observed statistically significant correlation among the two measures, albeit of a relatively low value of Pearson's r. These observations, taken together, strengthen our initial hypothesis of general GPCR-RAMP interaction. Specifically, when comparing the expression change of the 14 MERFISH-examined GPCRs upon RAMP2 over-expression as measured by RNA-seq and by the MERFISH method, we did not see any significant correlation among the two (Figure S5B, see Discussion).

DISCUSSION

In this study, we present supportive evidence for GPCR-RAMP interactions based on RNA expression patterns in HEK293T cells. We adapted a recently reported multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) approach and used it to count the number of candidate GPCR RNA transcripts per cell. The single-molecule technique that we utilized was based on the one presented by Chen et al. (Chen et al., 2015) and was modified to better suit efficient detection of a smaller number of genes and with fewer hybridization rounds. Chen et al. used 16 hybridization rounds to resolve 140 genes with error detection and correction, whereas we were able to resolve 14 genes with 8 hybridization rounds. Our adaptation to the multiplexed *in situ* hybridization protocol adds flexibility and shows that experiments can be designed in different ways to target greater or fewer numbers of molecules depending on the research question while minimizing the number of imaging rounds and the technical burden of synthesizing large probe libraries. Although our method is relatively simplified compared with Chen et al., we were able to reliably detect single genes with a lower number of fluorescent probes. Other studies have reported results using even fewer than 10 encoding probes (Batish et al., 2011; Buxbaum et al., 2014; Lubeck and Cai, 2012). The robust detection efficiency we observed can be attributed, at least in part, to the probe synthesis method we used. Unlike the bulk probe synthesis utilized by others (Chen et al., 2015), we used an individual, probe-by-probe synthesis, including a separate quality control step for each probe (see "Probe Design and Synthesis" in Transparent Methods). Our probe synthesis protocol also allowed us to control the concentration of each different probe, which is not generally possible with bulk synthesis.

The MERFISH results were consistent with the results obtained from our previous phylogenetic analysis (Barbash et al., 2017). As co-evolution of gene pairs could be an indicator for gene interaction, we expected to observe a change in the expression for GPCR transcripts upon RAMP2 over-expression in cases in which the GPCR co-evolved with RAMP2. Based on the previous literature on the mechanisms by which RAMP2 affects GPCRs (McLatchie et al., 1998; Parameswaran and Spielman, 2006), a change in GPCR expression is likely to be driven by a direct effect of RAMP2 on GPCR signaling or trafficking, which in turn could affect synthesis levels via feedback regulation, as was shown before (Black et al., 2016; Lefkowitz and Shenoy, 2005; Tsao et al., 2001). Indeed, we observed a correlation between expression effects on GPCRs upon over-expression of RAMP2 and the degree of co-evolution. The feedback regulation, together with the fact that we measured the association of phylogenetic patterns with expression, are probably the reasons for the relatively subtle effect we observed in the RNA *in situ* hybridization experiments (Figure 3). Nevertheless, the expression data generally support the approach of utilizing phylogenetic data to estimate gene pair interaction and specifically support GPCR-RAMP2 interactions for the receptors tested.

It is worth noting that some GPCRs showed larger expression changes in our mock transfection rather than RAMP2 transfection experiments. One possible explanation for this observation could be an opposing effect of RAMP2 and mock transfections. For example, if the expression of a given GPCR was up-regulated under the existence of a transfection construct, and the effect of RAMP2 over-expression on the GPCR would be to down-regulate it, the outcome could be larger expression change for the mock condition.

The five GPCRs that show a significant change of expression in MERFISH upon RAMP2 over-expression are involved in a variety of human diseases. ADORA2B is an adenosine GPCR, activating adenylatecyclase in the presence of a ligand. ADORA2B modulates inflammation, and as such is involved in inflammatory diseases (Hasko et al., 2008). OPRD1 is an opioid receptor, which reacts with the endogenous ligand enkephalin, is expressed in particularly high levels in the basal ganglia and neocortical regions of the brain (Peppin and Raffa, 2015), and is involved in major depressive disorder (Kempton et al., 2011). F2RL3 codes for coagulation factor II (thrombin) receptor-like 3, activated by proteases and is involved in cardiovascular disease (Leger et al., 2006). GPR141 has some evidence for involvement in autism spectrum disorders (Mitra et al., 2017). P2RY8 codes for P2Y purinoceptor 8 activated by adenosine and uridine nucleotides and is associated with diffuse large B cell lymphoma (Lohr et al., 2012). No previous receptor-RAMP interaction was reported before for any of these GPCRs, and further study is warranted to delineate the nature of this interaction and determine its involvement in the above-mentioned diseases.

We also examined our hypothesis of global GPCR-RAMP expression interaction using an independent method for expression profiling, i.e., RNA-seq. We observed a global shift in GPCR expression upon RAMP2 over-expression and global correlation with our previously reported phylogenetic measures. Specifically, comparing the MERFISH and RNA-seq methods, across the 14 GPCRs that overlap the two methods, we did not find significant expression correlation. This lack of correlation could be due to

different sources of technical noise in those methods. For example, whereas MERFISH is a single-cell, single-molecule method, RNA-seq is a whole-tissue method. In addition, each of the methods has intrinsic biases as to the transcripts that are more easily detected by it. For example, MERFISH strongly depends on RNA linearity, whereas RNA-seq strongly depends on the GC content. Importantly, however, although the two methods do not correlate on the specific set of 14 GPCRs, they both point at the same conclusion: that RAMP2 over-expression leads to a detectable change in GPCR population and that this change is in concordance with phylogenetic measures of GPCR-RAMP co-evolution.

We interpret the results of our over-expression experiments as supportive evidence for a global GPCR-RAMP interaction. It is worth mentioning that an alternative explanation could contribute to the observed results. Previous studies showed that over-expression of transmembrane proteins could indirectly change the mRNA levels of other transmembrane proteins (Buskirk and Green, 2017; Hollien and Weissman, 2006) with the following presumable mechanism. Typically, ribosomes translate the mRNA of transmembrane proteins until encountering the membrane translocation signal (signal sequence), which results in the recruitment of the signal recognition particle (SRP). The ribosome then docks to the translocon in the endoplasmic reticulum following by dissociation of the SRP before any further translation of the polypeptide can occur. Introducing high levels of a signal-sequence-harboring mRNA (e.g., a RAMP mRNA) can lead to competition on the limited number of translocons. This competition could lead to stalled translation, and as a result, to mRNA degradation of transmembrane proteins.

In summary, our results based on RNA expression analysis give further evidence for the existence of a global GPCR-RAMP interaction map. The correlation of message levels for GPCRs and RAMPs supports the hypothesis that GPCR-RAMP interactions are widespread and have functional consequences. As GPCR signaling is a fundamental mechanism across animal organisms, tissues, and cell types, further detailing of the GPCR-RAMP interaction map is warranted.

Limitations of the Study

We carried out the MERFISH analysis to evaluate the effect of RAMP2 over-expression on a limited number of 14 candidate GPCRs. Our analysis in the current study is limited almost entirely to the RNA level, which indirectly supports functional interactions at the protein level. However, future studies will be required to validate candidate GPCR-RAMP interactions and assess their functional consequences.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is Gene Expression Omnibus: GSE122633.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, five figures, two tables, and one data file and can be found with this article online at https://doi.org/10.1016/j.isci.2018.12.024.

ACKNOWLEDGMENTS

We acknowledge the generous support of the Rothschild post-doctoral fellowship (S.B.), the BengtWinblad's Foundation for Young Scientists, the Margareta af Ugglas Foundation, the Gun och Bertil Stohnes Foundation (T.P.), the Nicholson Short-Term Exchange fellowship (E.L.), the Robertson Therapeutic Development Fund, the Crowley Family Fund, and the Danica Foundation (T.H.). We also acknowledge the Rockefeller Bio-Imaging Resource Center (BIRC) and particularly Dr. Kaye Thomas for technical support and guidance.

AUTHOR CONTRIBUTIONS

S.B., T.P.S., T.H., E.L., and T.P. designed experiments and analyses; S.B., M.A.K., and T.P. performed experiments; S.B. performed computational analyses; and S.B., T.P.S., T.H., E.L., and T.P. wrote the paper.

DECLARATION OF INTERESTS

The authors have no financial interests to declare.

Received: January 31, 2018
Revised: September 4, 2018
Accepted: December 21, 2018
Published: January 25, 2019

REFERENCES

Atwood, B.K., Lopez, J., Wager-Miller, J., Mackie, K., and Straiker, A. (2011). Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. BMC Genomics *12*, 14.

Barbash, S., Lorenzen, E., Persson, T., Huber, T., and Sakmar, T.P. (2017). GPCRs globally coevolved with receptor activity-modifying proteins, RAMPs. Proc. Natl. Acad. Sci. U S A *114*, 12015–12020.

Batish, M., Raj, A., and Tyagi, S. (2011). Single molecule imaging of RNA in situ. Methods Mol. Biol. *714*, 3–13.

Black, J.B., Premont, R.T., and Daaka, Y. (2016). Feedback regulation of G protein-coupled receptor signaling by GRKs and arrestins. Semin.Cell Dev. Biol. *50*, 95–104.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20160183.

Buxbaum, A.R., Wu, B., and Singer, R.H. (2014). Single beta-actin mRNA detection in neurons reveals a mechanism for regulating its translatability. Science *343*, 419–422.

Chen, K.H., Boettinger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). RNA imaging.Spatially resolved, highly multiplexed RNA profiling in single cells. Science *348*, aaa6090.

Hasko, G., Linden, J., Cronstein, B., and Pacher, P. (2008). Adenosine receptors: therapeutic aspects for inflammatory and immune diseases. Nat. Rev. Drug Discov. *7*, 759–770.

Hay, D.L., and Pioszak, A.A. (2016). Receptor activity-modifying proteins (RAMPs): new insights and roles. Annu. Rev. Pharmacol. Toxicol. *56*, 469–487.

Hay, D.L., Poyner, D.R., and Sexton, P.M. (2006). GPCR modulation by RAMPs. Pharmacol.Ther. *109*, 173–197.

Hollien, J., and Weissman, J.S. (2006). Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. Science *313*, 104–107.

Kempton, M.J., Salvador, Z., Munafo, M.R., Geddes, J.R., Simmons, A., Frangou, S., and Williams, S.C. (2011). Structural neuroimaging studies in major depressive disorder.Meta-analysis and comparison with bipolar disorder. Arch. Gen. Psychiatry *68*, 675–690.

Lefkowitz, R.J., and Shenoy, S.K. (2005). Transduction of receptor signals by beta-arrestins. Science *308*, 512–517.

Leger, A.J., Covic, L., and Kuliopulos, A. (2006). Protease-activated receptors in cardiovascular diseases. Circulation *114*, 1070–1077.

Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnex, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., et al. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc. Natl. Acad. Sci. U S A *109*, 3879–3884.

Lubeck, E., and Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat. Methods *9*, 743–748.

McLatchie, L.M., Fraser, N.J., Main, M.J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M.G., and Foord, S.M. (1998). RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature *393*, 333–339.

Mitra, I., Lavillaureix, A., Yeh, E., Traglia, M., Tsang, K., Bearden, C.E., Rauen, K.A., and Weiss, L.A. (2017). Reverse pathway genetic approach identifies epistasis in autism spectrum disorders. PLoS Genet. *13*, e1006516.

Parameswaran, N., and Spielman, W.S. (2006). RAMPs: the past, present and future. Trends.Biochem.Sci. *31*, 631–638.

Peppin, J.F., and Raffa, R.B. (2015). Delta opioid agonists: a concise update on potential therapeutic applications. J. Clin. Pharm. Ther. *40*, 155–166.

Sharp, S.I., Hu, Y., Weymer, J.F., Rizig, M., McQuillan, A., Hunt, S.P., and Gurling, H.M. (2013). The effect of clozapine on mRNA expression for genes encoding G protein-coupled receptors and the protein components of clathrin-mediated endocytosis. Psychiatr.Genet. *23*, 153–162.

Tholanikunnel, B.G., and Malbon, C.C. (1997). A 20-nucleotide (A + U)-rich element of beta2-adrenergic receptor (beta2AR) mRNA mediates binding to beta2AR-binding protein and is obligate for agonist-induced destabilization of receptor mRNA. J. Biol. Chem. *272*, 11471–11478.

Tsao, P., Cao, T., and Von Zastrow, M. (2001). Role of endocytosis in mediating downregulation of G-protein-coupled receptors. Trends.Pharmacol. Sci. *22*, 91–96.

Venkatakrishnan, A.J., Depui, X., Lebon, G., Tate, C.G., Schertler, G.F., and Babu, M.M. (2013). Molecular signatures of G-protein-coupled receptors. Nature *494*, 185–194.

Weston, C., Lu, J., Li, N., Barkan, K., Richards, G.O., Roberts, D.J., Skerry, T.M., Poyner, D., Pardamwar, M., Reynolds, C.A., et al. (2015). Modulation of glucagon receptor pharmacology by receptor activity-modifying protein-2 (RAMP2). J. Biol. Chem. *290*, 23009–23022.

ISCI, Volume *11*

Supplemental Information

Detection of Concordance between Transcriptional Levels of GPCRs and Receptor-Activity-Modifying Proteins

Shahar Barbash, Torbjörn Persson, Emily Lorenzen, Manija A. Kazmi, Thomas Huber, and Thomas P. Sakmar

Supplementary Information

Detection of Concordance between Transcriptional Levels of GPCRs and Receptor Activity-Modifying Proteins

Supplementary Figures

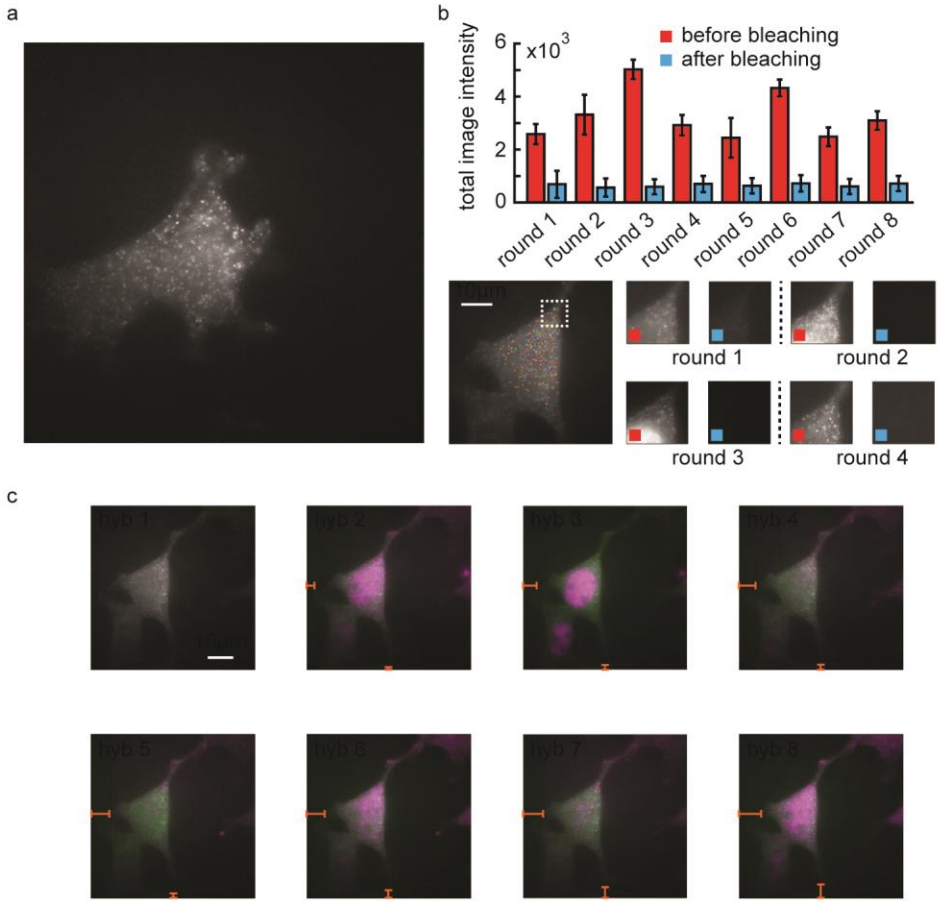


Figure S1: MERFISH bleaching efficiency and image registration, related to Figure 1. (a) Example of a full field raw image for one cell in one hybridization round. (b) Bleaching efficiency between hybridization rounds. Bar graph (mean, s.e.m) for total image intensity of unfiltered images for before and after bleaching procedure (see Methods for details). Also shown are zoomed-in (dashed white square in image) example images for the first four hybridization rounds. (c) Image registration (intensity based, 'rigid') based on beads detection (not shown). Shown are image composites of the first hybridization image (green) with each of the other images (purple) and the shifts in the X and Y directions as orange bars.

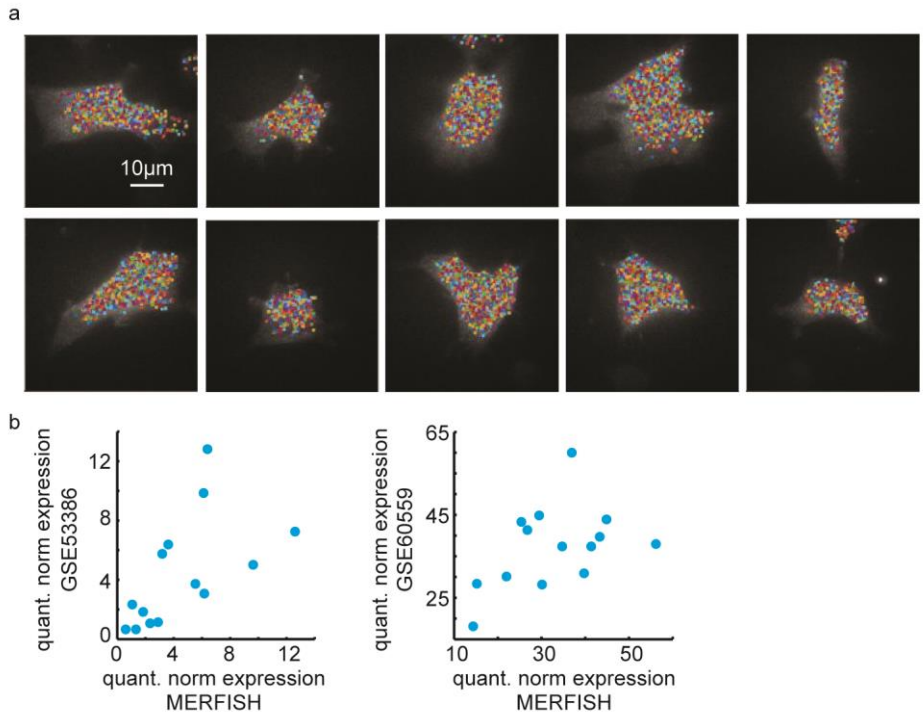


Figure S2: Example expression in cells and correlation with available datasets, related to Figure 1. (a) Ten example cells overlaid with the 14 detected GPCRs. (b) Scatter plots comparing our GPCR expression data with Control HEK293T samples from the GEO submitted datasets GSE53386 and GSE60559 (Pearson`s *r*=0.64, *P*=0.056 for GSE53386 and *r*=0.72, *P*=0.040 for GSE60559). Data are quantile normalized (quant. norm.) between MERFISH and each of the datasets separately.

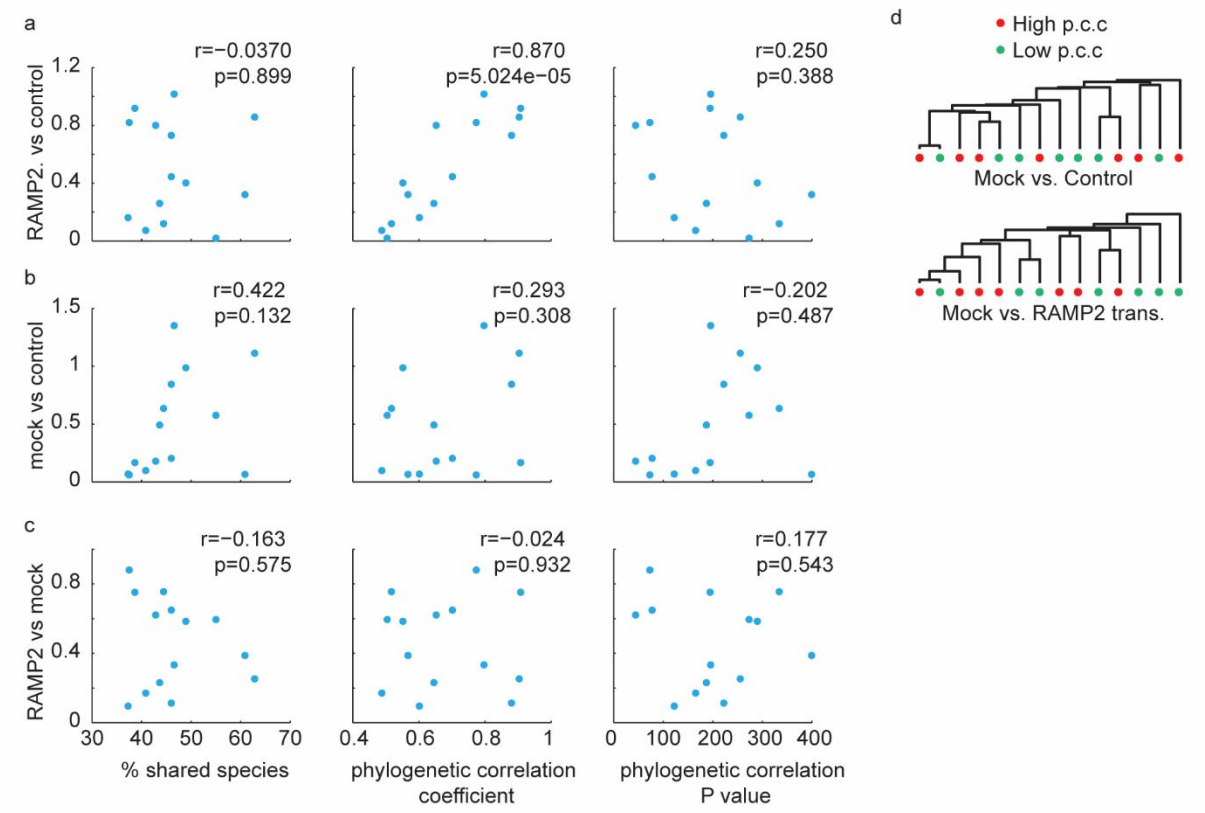


Figure S3: No correlation between phylogenetic measures and Mock transfection-induced expression change, related to Figure 3. (a) Scatter plot of absolute value of log2-fold change of GPCR expression between HEK293T cells RAMP2-transfected and control cells against %-shared species, phylogenetic correlation coefficient and P value. For each plot Pearson`s *r* and p values are shown. (b) Similar scatter plots for Mock-transfection *versus* control. (c) Similar scatter plots for RAMP2-transfection *versus* Mock-transfection. (d) Expression-based dendrograms, calculated as in Figure 3c, for Mock transfection *versus* no transfection (Control) and Mock transfection *versus* RAMP2 transfection.

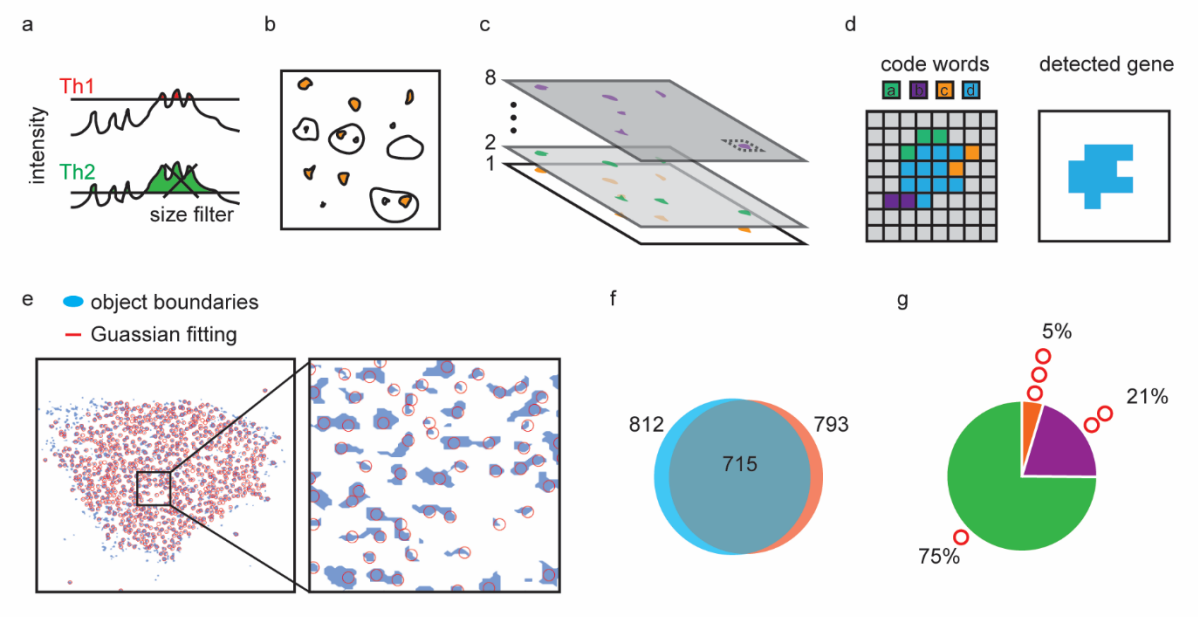


Figure S4: Image analysis pipeline, related to Figures 1 and 3. (a) Transforming the intensity images to binary images by thresholding at different intensity levels. (b) After size filtration of identified objects, objects of between 10-50 pixels were left. (c) Code word construction across the eight images per imaging field. (d) Identical neighboring code words clusterization and gene detection based on a matching code word for any of the 14 GPCRs, or a 1 bit correction based on the smallest Hamming Distance. (e) Object boundaries method for detecting fluorescent signals (as described here; blue) and Gaussian fitting method (red) overlaid for an example cell. A zoom-in of the marked area is also shown. (f) Venn diagram for the number of spots detected by the 'object boundaries' method (blue) and the Gaussian fitting (red) method and the overlapping identified objects for the cell shown in e. (g) Pie chart for the identified objects from both methods (shared area in f) showing the number of Gaussians fitted inside a single object detected by the 'object boundaries' method (one, two or three Gaussians). That in some cases more than one Gaussian is fitted for a single region from the 'object boundaries' method is also seen in the zoom-in area in e. This analysis was performed for 10 imaging fields in total and showed similar results.

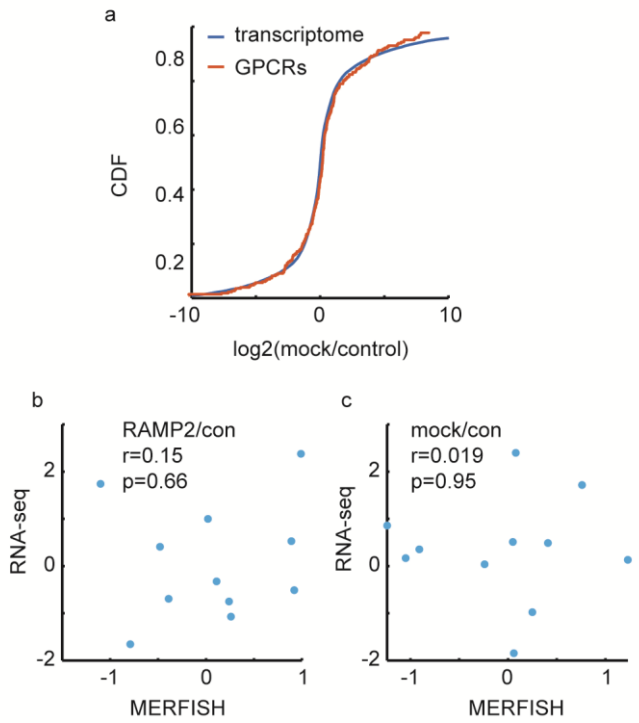


Figure S5: Global expression effect of RAMP2 over-expression on GPCRs as measured with RNA sequencing, related to Figure 4. (a) Cumulative Distribution Function (CDF) of log2 of fold change between the average across three samples of mock-transfection and three control samples. CDF is shown separately for all transcriptome (red) and for all GPCR transcripts (red). Kolmogorov-Smirnov P value for difference between distributions: 0.27. (b) Scatter plot of the log2 fold change for RAMP2 versus control (i.e. untransfected cells) as measured by RNA sequencing (Y axis) and MERFISH (X axis). Also shown are Pearson`s r and P value for the correlation between the two measures. (c) Scatter plot of the log2 fold change for mock-transfection versus control as measured by RNA sequencing (Y axis) and MERFISH (X axis).

Supplementary Tables

Table S1: 14 GPCRs detected by multiple *in situ* hybridization, related to Figure 1. Shown are their official gene name, their RefSeq annotation and the code word used for detection.

	Gene	RefSeq	Word Code							
1	ADORA2B	NM_000676	0	0	1	0	1	1	0	1
2	S1PR1	NM_001400	0	1	0	1	1	1	0	0
3	NTSR1	NM_002531	1	0	0	0	1	1	1	0
4	OPRD1	NM_000911	0	1	0	0	1	0	1	1
5	F2RL1	NM_005242	0	1	1	0	0	1	1	0
6	F2RL3	NM_003950	0	0	0	1	0	1	1	1
7	GPR50	NM_004224	1	1	1	0	1	0	0	0
8	GPR141	NM_181791	1	0	0	1	1	0	0	1
9	GPR160	NM_014373	0	0	1	1	1	0	1	0
10	GPR176	NM_007223	1	1	0	0	0	1	0	1
11	LGR4	NM_018490	1	0	1	1	0	1	0	0
12	P2RY8	NM_178129	1	0	1	0	0	0	1	1
13	CELSR2	NM_001408	1	1	0	1	0	0	1	0
14	FZD1	NM_003505	0	1	1	1	0	0	0	1

Table S2: Fluorescent probes used to detect genes, related to Figure 1. These 8 probes sequences are taken from Chen *et al.* (Chen et al., 2015) and they hybridize with the flanking tails of the gene probes in Table S8. All had 3`-end modification of an added Cy5 fluorescent dye.

	Probe sequence
1	CGCAACGCTTGGGACGGTTCCAATCGGATC
2	CGAATGCTCTGGCCTCGAACGAACGATAGC
3	ACAAATCCGACCAGATCGGACGATCATGGG
4	CAAGTATGCAGCGCGATTGACCGTCTCGTT
5	AAGTCGTACGCCGATGCGCAGCAATTCACT
6	CGAAACATCGGCCACGGTCCCGTTGAACTT
7	GCATGAGTTGCCTGGCGTTGCGACGACTAA
8	GGCCAATGGCCCAGGTCCGTCACGCAATTT

Table S3: Cell medium and buffers used for RNA *in situ* hybridization experiments, related to Figures 1, 2 and 3. All buffers were prepared in Nuclease-free water (Ambion) unless otherwise stated.

Cell medium	DMEM with Glutamax™ + 10% FBS + 1% PenStrep (ThermoFisher Scientific)
Fixation Buffer	4% Formaldehyde (Thermo Fisher Scientific) in PBS
Encoding Wash Buffer	2x Saline-Sodium Citrate; SSC, 30% Formamide (Ambion), 2 mM Vanadyl Ribonucleoside complex (VRC)(New England Biolabs).
Encoding Probe Solution	10 % (w/v) dextran sulfate, 1 mg/ml yeast tRNA (Sigma-Aldrich), 2x SSC, 30% Formamide, 2 mM VRC, 6.25 pM of each of the 140 encoding probes.
SSC buffer	2x SSC (0.3 M NaCl, 0.03 M Sodium Citrate)
Readout Probe Solution (8 different solutions, one for each readout probe)	10 % (w/v) dextran sulfate, 2x SSC, 10% Formamide, 2 mM VRC, 100 pM readout probes.
Readout Wash Buffer	2x SSC, 20% Formamide, 2 mM VRC
Trolox solution	10 % (w/v) (±)-6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox)(Sigma-Aldrich) in methanol
Imaging Buffer	2x SSC, 50 mM Tris pH 8.0 (Ambion), 10% D-(+)-Glucose (Sigma-Aldrich), 0.1% (w/v) Glucose Oxidase (Sigma-Aldrich), 0.5% (v/v) Trolox solution, 200 µg/ml Catalase (Sigma-Aldrich)
Bleaching buffer	2x SSC, 2 mM VRC

Datasets

Dataset 1: Full RNA-targeting probe sequences used in the study, related to Figure 1. Shown are the ten probes used for each of the 14 GPCRs. Upper case denotes flanking tail parts and lower case denotes RNA targeting region. Attached as a separate excel file.

Transparent Methods

Criteria of probe design and synthesis

We designed and implemented a modified version of the *in situ* hybridization method described by Chen *et al.* (Chen et al., 2015). For reasons of simplicity and to streamline the experimental protocol, we performed eight hybridization rounds compared with 16 as presented by Chen *et al.* In each hybridization round fluorescent signals are detected, and across the eight rounds a binary 'code word' was constructed for each pixel in the image. In the

code word, a '1' denotes detection in a given location, and a '0' indicates no detection. Therefore, each code word is a binary representation of the series of detections ('hit' or 'no-hit'), of an RNA molecule in an image. Similarly to Chen *et al.* we designed the code words to have Hamming Distance 4. The Hamming Distance is the number of permutations needed to be done to change one 'code word' to another. Hamming Distance 4 allows error detection and correction in the code word decoding stage. When a detection of a code word that is different in one bit from an expected code word occurs, Hamming Distance 4 allows for correction to the correct code word (*i.e.,* the correct gene). Hamming Distance 4 ensures that a distance of a single bit from a code word could only be for one measured gene and hence enables correction for that gene. For more details about Hamming Distance 4 and error detection and correction see Chen *et al.* (Chen et al., 2015).

The maximal number of genes that could be detected with eight hybridizations and a Hamming Distance 4, are 14. We compiled a list of 14 GPCRs that had the following characteristics. First, their transcript length enabled targeting by ten or more proper probes (see criteria below) of 30-nt with a 30-nt gap between probes. Second, these GPCRs were found to be expressed in HEK293T cells previously (Supplementary Fig. 2b). Third, these GPCRs have a homogeneous distribution across the calculated phylogenetic correlation coefficient (seen as distribution along the X axis in Supplementary Fig. 3a). This homogeneous distribution is important in order to represent both high and low RAMP2 phylogenetic associations.

Probes were designed using the following criteria: melting temperature of 65°C to 80°C, G-C content between 45% and 55%, absence of stable hairpins and absence of dimerization with other probes in the experiment. For specificity, the probes were screened against the entire human transcriptome with BLAST using the mature RNA list from Ensemble (updated for December 2016), and any nonspecific hits were filtered out. Next, for each GPCR ten probes were chosen so that their localization distributed along the RNA molecule would be homogenous as possible in order to represent all parts of the transcript. The number of probes was set to ten per gene to allow for comparable fluorescent intensity across genes in an image. These probes were concatenated on both sides with flanking ends of fluorescent hybridizing regions (*i.e.,* readout sequences; Supplementary Table 2) to give 140 final detection probes (see Dataset 1 for full sequences). The readout sequences were assigned to the encoding probes so that each readout sequence will be localized in a homogeneous manner along the RNA species and will have the same frequency as other readout sequences. All probes were synthesized in separate reactions by Integrated DNA Technologies (IDT). Note that the endogenous RAMP2 transcript length is 808-nt, which is too short for detection with the current protocol after allowing for a 30-nt gap between probe positions and filtering by the probe criteria mentioned above. Fluorescent probes were designed to hybridize with the part of the encoding probes that do not hybridize with the target RNA. Eight such probes were adopted from Chen *et al.*, synthesized and modified by adding a Cy5 fluorescent dye on their 3' ends (by IDT).

Sample preparation

The protocol for sample preparation was adapted from Chen *et al.* (Chen et al., 2015) with a few modifications. The details for the cell medium and buffers used can be found in Supplementary Table 3. HEK293T cells were cultured in cell medium and 300,000 cells were

seeded onto poly-D-lysine mol wt 70,000-150,000 (Sigma-Aldrich) coated cover glasses (18-mm, thickness No. 1.5H, Marienfeld Superior) placed in 60-mm tissue culture dishes. The next day, cells were transfected with 4 µg plasmid DNA for MERFISH, or 2 µg plasmid DNA for RNA sequencing, using Lipofectamine 2000 (Thermo Fisher Scientific) according to manufacturer's protocol. Plasmid DNA included RAMP2 with a FLAG-tag epitope sequence following the signal peptide, OLLAS-epitope (Rico et al., submitted) tag on the C-terminus in pcDNA3.1, codon optimized for expression in human cell lines. After 48 h incubation at 37°C, RNA was extracted using the miRneasy kit (QIAGEN) for RNA sequencing, or separately, samples for MERFISH were added a fixation buffer. For the MERFISH samples, after 20 min incubation at room temperature, the fixation buffer was removed, and the samples were incubated for 5 min at room temperature with 0.1% (w/v) sodium borohydride (Sigma-Aldrich) diluted in RNAse free water. Next, cells were washed three times with ice-cold PBS and permeabilized for 2 min at room temperature using 0.5% (v/v) Triton (Sigma-Aldrich). The samples were then washed three times again with ice-cold PBS followed by 5 min incubation with Encoding Wash Buffer at room temperature. Next, 4 ml Encoding Probe Buffer was added to the cells and the samples were incubated at 37°C overnight. The next day, Encoding Wash Buffer heated to 47°C was added to the cells. The samples were incubated at 47°C for 10 min. This step was repeated for two additional times. Next, Carboxylate-Modified Microspheres, 0.2 µm (540/560) diluted 1:20,000 in SSC buffer were sonicated for 90 sec and then added to the cells. The samples were incubated with beads for 5 min at room temperature and then washed once with SSC buffer. This wash was followed by another incubation with fixation buffer for 30 min at room temperature. Before starting the imaging procedure, cells were washed three times in 2x SSC buffer. The samples were covered from light and kept in SSC buffer until imaged. Samples were not stored for longer than one week.

Imaging

For detailed sample preparation see Supplementary Experimental Procedures. Samples were placed inside an RC-30WA flow chamber (Warner Instruments), and the flow through the chamber was controlled manually with appropriate tubing and syringes. Imaging was done on a Nikon TiE inverted microscope with oblique incidence excitation, 100x oil immersion objective, and Andor Neo sCMOS camera. A pixel corresponded to ~100nm in the sample plane. The Elements Software was used to control the microscope and capture images. We started each imaging session by flowing 2 ml of Bleaching Buffer through the sample. We then chose 20-30 cells of a relatively flat morphology and reduced auto-fluorescent noise by bleaching each cell with a 647-nm laser of 100% intensity for 1 sec. Then, 1 ml of fluorescent probe buffer with the first fluorescent probe was flowed in and incubated for 15 min. Fluorescent probe buffer was washed with 2 ml of the Washing Buffer following another incubating time of 3 min. Next, 2 ml of the Imaging Buffer was flowed in. At this point, the total-internal reflectance fluorescence (TIRF) angle was optimized for the sample, and the cells were captured with bright field: 561 nm (for nano-beads detection) and 647 nm consecutively in an automated manner with the Perfect Focus mechanism activated. Fresh 2ml of Bleaching Buffer were immediately flowed in, and the cells were bleached with the 647-nm laser at 100% intensity for 1 sec. For bleaching efficiency analysis the post-bleach cells were captured again with 647-nm excitation. Then 1ml fluorescent probe buffer with the second fluorescent probe was flowed in and the whole cycle

was repeated. This cycle was done consecutively eight times, each time with a different fluorescent probe, in the order shown in Supplementary Table 2. Note that for the shuffled code book validation experiment the probes order was changed (Fig. 1f). In addition to the shuffled code book validation experiment, we performed two more validations. First, to evaluate autofluorescence, we imaged the sample using only encoding probes. Second, to evaluate the degree of nonspecific binding, we imaged the sample using only readout probes. In both cases, the overall intensity was less than 5% of that of the averaged intensity across the eight hybridization rounds for the full experiment. See Supplementary Experimental Procedures for full description of image analysis.

Image analysis

In our image analysis approach, we adopted elements from the analysis reported by Chen *et al.* and elements from the analysis reported by Moffitt *et al.* (Moffitt et al., 2016). For each cell in each imaging session, all eight images were registered, based on the signal from the Carboxylate-Modified Microspheres (nano-beads), against the first image (that of hybridization round 1) with intensity-based, rigid transformation consisting of translation. (See Supplementary Figure 1c for an example registration.) Registration was done using the Image Processing Toolbox of Matlab2017b. Fluorescent spots were detected by transforming the registered images to binary images by thresholding at different intensity levels in the 70% top range of intensities. This thresholding was done to allow for objects of different intensities in different regions of a cell, as well as between cells and between experiments, to be detected. Next, the boundaries of the objects in the binary images were traced by the Moore-Neighbor tracing algorithm and only objects of 10-50 pixels were retained. This range is compatible with the size of objects identified by others (Chen et al., 2015). Next, for each location (pixel) across the eight images of a cell a code word was constructed, where a '1' denotes signal detection and a '0' denotes no detection. Identical neighboring code words were clustered together to build a new object, and these objects were again filtered for size (10-50 pixels retained). Next, the Hamming distance of each of these code words with each of the 14 GPCR code words was calculated. Exact matches (Hamming Distance of 0) were defined as a detected gene, and Hamming Distance of 1 was corrected to the closest gene's code word (only one such word exists as per the design). These analysis steps are presented in Supplementary Figure 4a-d. We compared the results of our detection protocol with Gaussian fitting results of the ThunderStorm ImageJ plugin tool (Ovesny et al., 2014) (http://zitmen.github.io/thunderstorm/) and found that ~90% of the fluorescence signals identified in each of the methods were identified by both methods (Supplementary Fig. 4e and f). In 75% of the cases a single Gaussian was detected per region from our 'object boundaries' method and in the remaining 25% two or three Gaussians were detected (Supplementary Fig. 4g). We decided to process images by the protocol described here due to lower computational burden and because we saw no biological reason to filter for round signals (circular or elliptical) which is a characteristic of Gaussian fitting. During image analysis, cells were filtered out based on the following criteria. Cells with no detected nano-beads in their image field, and hence to which no registration could be performed, as well as cells that had beads, but to which registration was unsuccessful, were filtered out. Cell images of contrast lower than 0.5 were filtered out. The contrast was

calculated as [max(p)-min(p)]/max(p), where p is pixels intensities of an image. Lastly, the 25% of cells with the lowest global detected RNA expression were filtered out.

Western Immunoblot Analysis

Cells were collected by repeatedly flowing PBS over the cells using the pipette until all cells were detached. The cells were then pelleted by centrifugation at 5,000 x g, dissolved in lysis buffer containing 50 mM Tris Base, 100 mM NaCl, 1 mM $CaCl_2$, 1x Protease inhibitor cocktail (Roche), 0.1% (v/v) PMSF, 1% (w/v) *n*-dodecyl-β-D-maltoside detergent and incubated with gentle rotation at 4°C for 1 hr. Next, the samples were centrifuged at 15,000 x g for 15 min, and the supernatant fraction was collected. Protein concentration was measured using DC Protein Assay (BioRad), and 20 µg of protein was mixed with 0.1 M DTT and NuPAGE LDS Sample Buffer (Thermo Fisher Scientific). The samples were heated at 70°C for 10 min before being separated on a NuPAGE 4–12% Bis-Tris gel (Thermo Fisher Scientific) and transferred to a PVDF membrane at 18V for 45 min using a Transblot SD Semidry Transfer Cell (BioRad). The membrane was blocked in a 1:1 mixture of PBS and Odyssey Blocking Buffer (LI-COR). Next, the membrane was incubated overnight at 4°C with primary antibodies (rabbit anti-Actin 1:10,000 and mouse anti-FLAG M2 1:2,000) diluted in blocking buffer. Next day, the membrane was washed 3 x 10 min in TBS-T followed by incubation with secondary antibodies (Goat anti-mouse IRDye 680 and goat anti-rabbit IRDye 800 (LI-COR)) diluted 1:20,000 in blocking buffer with 0.1% (v/v) Tween and 0.01% (w/v) SDS. This incubation was followed by another 3 x 10 min wash in TBS-T and a final wash in PBS before detecting the signal in a LiCor Odyssey SA instrument.

RNA Sequencing

See samples preparation in Supplementary experimental procedures. Samples were examined with Bioanalyzer for RNA integrity and concentration before sequencing. RNA integrity number was 9.62 with 0.98 standard deviation. 100 ng of total RNA was used to generate RNA-Seq libraries using Illumina TruSeq stranded mRNA LT kit (Cat# RS-122-2101). Libraries prepared with unique barcodes were pooled at equal molar ratios. The pool was denatured and sequenced on Illumina NextSeq 500 sequencer using high output V2 reagents and NextSeq Control Software v1.4 to generate 75 bp single reads, following manufactures protocol (Cat# 15048776 Rev.E). The full sequencing data (mapped read and FPKM files) are available at Gene Expression Omnibus, accession number GSE122633. For read alignment after adaptor cutting we used STAR software. Quality control was verified with qualilmap. Average number of input reads to the read aligner, across the nine samples was 40 million reads with 1 million standard deviation and the percentage of uniquely mapped reads was 88.7% with 2.6% standard deviation. For FPKM calculation we used cufflinks.

Code availability

Code used to conduct the data analysis is available upon request from sbarbash@rockefeller.edu.