GPCR-G protein selectivity – a unified meta-analysis

Alexander S. Hauser[1], Charlotte Avet[2], Claire Normand[3], Arturo Mancini[3], Asuka Inoue[4], Michel Bouvier[2*] and David E. Gloriam[1*]

[1] Department of Drug Design and Pharmacology, University of Copenhagen, 2100 Copenhagen, Denmark.

[2] Institute for Research in Immunology and Cancer (IRIC), and Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, Québec, H3T 1J4, Canada.

[3] Domain Therapeutics North America, Montréal, Québec, H4S 1Z9, Canada.

[4] Graduate School of Pharmaceutical Sciences, Tohoku University, Sendai, Miyagi 980-8578, Japan.

* Corresponding authors:

Data analysis: David E. Gloriam, david.gloriam@sund.ku.dk

Pharmacology: Michel Bouvier, michel.bouvier@umontreal.ca

Abstract

Two-thirds of human hormones and one-third of clinical drugs act on membrane receptors that couple to G proteins to achieve appropriate functional responses. While G protein transducers from literature are annotated in the Guide to Pharmacology database, two recent large-scale datasets now expand the receptor-G protein 'couplome'. However, these three datasets differ in scope and reported G protein couplings giving different coverage and conclusions on GPCR-G protein signaling. Here, we report a meta-analysis unifying GPCR-G protein coupling, by standardized normalization and consensus support, into a common coupling map. This unravels novel consensus couplings for receptors supported by two independent sources and insights on coupling selectivity of GPCRs and classification of co-coupling G proteins. The coupling protocol, map and selectivity resources will promote advances in receptor research and cellular signaling towards the exploitation of G protein signaling specificity in design of safer drugs.

Introduction

G protein-coupled receptors (GPCRs) represent the largest family of proteins involved in signal propagation across biological membranes. They recognize a vast diversity of signals going from photons and odors to neurotransmitters, hormones, and cytokines (1). Their main signaling modality involves the engagement and activation of G proteins. G proteins are heterotrimeric proteins consisting of a α, β and γ subunits that dissociate to α and a βγ upon activation by a GPCR. G proteins are named by their α subunit (16 in human) and are divided into four families which share homology and downstream signaling pathways: G_s (G_s and G_{olf}), $G_{i/o}$ (G_{i1}, G_{i2}, G_{i3}, G_o, G_z, G_{t1}, G_{t2} and G_{gust}), $G_{q/11}$ (G_q, G_{11}, G_{14} and G_{15}) and $G_{12/13}$ (G_{12} and G_{13}). A GPCR's activation of G proteins can be very selective or promiscuous and change upon ligand-dependent biased signaling that alters its profile on the G protein subtype or family levels. The pleiotropic signaling and ligand-dependent bias of GPCRs pose a grand challenge in human biology to map the differential activation of specific G proteins.

The IUPHAR/BPS Guide to Pharmacology (GtP) database contains reference data from expert curation of literature (1). GtP couplings covers 253 GPCRs and the four G protein families. The G protein families have been classified into "primary" and "secondary" transducers without quantitative values. Recently, the Inoue group determined the first large-scale quantitative coupling profiles of 148 GPCRs to the G_q wildtype and 10 G protein chimeras employing a TGF-α shedding assay (2) (NTS_1 and TRH_1 added herein making it 150 receptors). Those chimeras consist of a G_q backbone in which the six most C-terminal Gα residues – a part of the H5 domain inserting to the intracellular receptor cavity – have been replaced to represent all 16 human G proteins (five of which have identical sequences to other G proteins, see below). In a paper accompanying the present analysis , the Bouvier group quantified the couplings of 100 GPCRs to 12 G proteins: G_s, G_{i1}, G_{i2}, G_o (G_{oA} and G_{oB} isoforms), G_z, G_q, G_{11}, G_{14}, G_{15}, G_{12} and G_{13} but not G_{olf} (couples mainly to olfactory receptors), G_{i3} and G_{t1-2} (Transducin, couples to rhodopsin (visual) receptors) and G_{gust} (Gustducin, couples to taste receptors) (3). The authors used novel enhanced bystander bioluminescence resonance energy transfer biosensors that allow to monitor G protein activation (G protein Effector Membrane Translocation assay; GEMTA) without need to modify the G protein subunits (except for G_s) or the receptors.

Here, we analyze the GtP, Inoue and Bouvier coupling datasets to determine confident couplings supported by at least two independent sources, including novel couplings discovered jointly by the two latter sources. We establish a scalable protocol to normalize quantitative G protein couplings, combine E_{max} and EC_{50} into a common $\log(E_{max}/EC_{50})$ (4) value and aggregate subtypes to allow comparisons across G protein families. On this basis, we develop a unified map of GPCR-G protein couplings that is also made available in the GPCRdb database hub (5), describe GPCR-G protein selectivity across an unprecedented number of receptors and coupling data points and reveal correlated co-couplings.

Results

Each profiling study doubled the receptors' average number of G protein family couplings

To obtain an overview of the coverages of GPCR-G protein coupling sources, we compared all couplings reported by both the Bouvier (3) and Inoue (2) groups and annotated in the Guide to Pharmacology database (GtP (1)) (Fig. 1a). This shows that the three sources together comprise couplings for 265 (67%) out of the 398 non-olfactory GPCRs and that 70 of these receptors are present in all datasets. The Bouvier and Inoue datasets have collectively quantified individual G protein couplings of 178 receptors using one assay, whereas the remaining 87 receptors have so far only been annotated in GtP on the G protein family level from a multitude of publications and assays (Fig. 1a).

To allow the comparison of coupling densities and distributions across datasets, we selected the E_{max} threshold (1.4 standard deviations above basal) that gives the best agreement between the Bouvier and Inoue dataset. We believe that this is the best possible means to estimate what is correct data (rather than false negative/positive couplings), as large-scale information about what G proteins and GPCRs coupling in nature is not available. This cut-off is more stringent the minimum of 3% signal above basal used in Inoue's original report (2). We also aggregated G protein subtype couplings of the families (see Methods). This reveals that while GtP covers the largest number of receptors they have relatively few couplings – 62% of all GPCR-G protein family pairs are 'non-couplers' (or not tested or annotated) compared to 35% in the Bouvier and 25% in the Inoue dataset (average of 1.5, 2.6 and 3.0 G protein families per GPCR, respectively; Fig. 1b). In particular $G_{12/13}$ couplings are underrepresented in GtP where they account for 3% of GPCR-G protein pair datapoints compared to 13% in Bouvier and 17% in Inoue. The Bouvier and Inoue datasets share a more similar overall distribution of couplings, expect for G_s coupling which is twice as frequent within the latter dataset (16% compared to 8%). This demonstrates that the two first systematic coupling profiling studies have substantially expanded the known GPCR-G protein 'couplome' and that their assay platforms are amenable to high-throughput profiling also for the $G_{12/13}$ family for which robust activation assays appeared only recently (6-8).

The Bouvier and Inoue datasets converge for 68% of G protein couplings, including 101 that are novel

Given that the Bouvier and Inoue groups used different biosensors and GtP annotates literature reports from very diverse assays, we sought to determine to which extent they report the same couplings for the same GPCRs i.e., their 70 common receptors (data in tab 'BIG-QualComp' in Spreadsheet S3). We find that all three sources agree on 39 (55%) of G_s, 45 (64%) of $G_{i/o}$, 35 (50%) of $G_{q/11}$ and 18 (25%) of $G_{12/13}$ couplings/non-couplings (mean 34 (49%), Fig. 2a) showing that the agreement varies substantially across the individual G protein families. When instead analyzing the agreement of just the quantitative studies (excluding GtP) and the G protein subtype level (Fig. 2b), it increases to an overall of 68% and ranging from at least 53% (for G_{15}) up to 81% (for G_{oA}). These findings define a sizeable reference set of consensus G protein couplings and show that consistent large-scale profiling studies generate more comparable results than literature.

We next identified the 'novel' G protein couplings for which a family annotation is missing in GtP but have high confidence from dual support by the Bouvier and Inoue groups (black in Fig. 2b). This revealed 38 receptors with novel couplings to 101 G proteins distributed across all families: G_s,:4, $G_{i/o}$:15, $G_{q/11}$:10 and $G_{12/13}$:21 (Fig. S1). The largest expansions – an increase by three of G protein families – was obtained for the histamine H_1 and endothelin ET_A receptors which was found to couple to all G protein families but only have $G_{q/11}$-coupling in GtP. Whereas it could be expected that GtP would miss couplings, we also analyzed if its expert curation excluded couplings that may be false positives as they are contradicted by both quantitative studies. This uncovered such G_s-coupling to the α_{2C}-adrenoceptor and cannabinoid CB_{1-2} receptors, $G_{12/13}$ coupling to the purinergic $P2Y_2$ receptor and $G_{i/o}$-coupling to the β_2-adrenoceptor, which however had weak G_z and G_{oB} coupling in the Bouvier study but did not cross the signal threshold (3). Notably, this is only 2% (5/254) of all GtP's GPCR-G protein family pairs. Taken together, these findings serve to quantify the expansion of the GPCR-G protein 'couplome' while also confirming the outstanding accuracy of the expert annotation in the GtP database. Therefore, the large number of new couplings results mainly from that these couplings were not investigated in literature and not inadequate curation.

Comparison across datasets reveal unsupported couplings requiring additional studies

While the Bouvier and recent Inoue datasets give an unprecedented mapping of GPCR activation on the single G protein subtype level, the above analyses point to a need to explain their observed differences to establish which couplings to use where datasets do not agree (Fig. 2). To this end, we analyzed the 70 common receptors for 'unsupported' and 'missing' couplings defined as being present or absent in only one of the three datasets (i.e., non-couplings and couplings, respectively, that are proposed by one but not supported by two datasets). For GtP, we consider coupling to a G protein subtype possible if a coupling has been observed from the respective family.

For G_{15}, 30% of Bouvier's couplings are unsupported (Fig. 3a). The Bouvier dataset has a large fraction, 37% of G_{15} couplings that are new compared to GtP whereas the Inoue dataset has no new couplings and is missing 13% G_{15} reported in GtP. This indicates that G_{15} couplings are underrepresented in both literature (which is annotated in GtP) as well as in Inoue's data. The few known couplings for G_{15} is likely explained by its lack of expression in the cells most used for *in vitro* experiments, such as HEK293 cells (3) and by the lack or weak (at very high concentration) effect of the $G_{q/11}$ inhibitor tools YM-254890 and FR900359, respectively (9). Furthermore, several of these novel G15 couplings were validated using Ca^{2+} assays in the accompanying paper from the Bouvier group (3). Taken together, these findings suggest that most of the unsupported G_{15} couplings in the Bouvier dataset are real (true positive) unique couplings.

For G_s and G_{12}, 27% and 24%, respectively of Inoue's couplings are unsupported (not present in Bouvier or GtP) and 29% of G_s and 49% of G_{12} couplings are new compared to GtP (Fig. 3b). In contrast, the Bouvier dataset has no and 13% unsupported G_s and G_{12} couplings, respectively and is missing 7% of G_s and 6% of G_{12} couplings in GtP. The lower representation of these two G proteins in Bouvier compared to Inoue may be due to the lower assay window (tab 'DataStats' in Spreadsheet S1). Therefore, additional studies would be needed to distinguish the real new G_s couplings in the Inoue dataset from non-couplers in GtP, which are likely many since robust assays and selective inhibitors have long been available for G_s signaling. Furthermore, the Bouvier data contains fewer G_{12} couplings than the Inoue group (32% relative 50%, overlap is 26%) that are new compared to GtP. Further studies with alternative assays maintaining a high G_{12} sensitivity would be desirable to gain independent support to these couplings.

For G_q, G_{11} and G_{14}, 13%, 16% and 16%, respectively of Inoue's couplings are unsupported (not present in Bouvier or GtP). The Bouvier dataset is instead missing 10%, 10% and 7%, respectively of the known couplings in GtP (Fig. 3b) and has only 1, 0 and 1 unique such couplings, respectively. This could be due to the fact that in many studies, people have equated Calcium signaling with G_q activation underestimating the role of G_i derived bg in the activation of the PLC-IP3-Calcium signaling pathway (10-13). Like for G_s and G_{15} above, further studies would be needed to reach consensus on whether these couplings are closest to the level of known in GtP or reported by Inoue group. Notably, although Inoue used the wildtype protein for G_q it has a similar or higher frequency of unsupported couplings compared to all other measured G proteins (G_q: 13% vs. G_{i1}: 7%, G_{i2}: 9%, G_{oA}: 7%, G_{oB}: 6%,

G_z: 13%, G_{11}: 16%, G_{14}: 16%, G_{15}: and G_{13}: 17%), except G_s and G_{12} (addressed above). In all, this suggests that the use of chimeric G proteins carrying over of G_q coupling to the other G proteins is not a main cause of Inoue's unsupported couplings.

To explain the reason to the unsupported couplings, beyond those addressed above for G_{15}, G_s and G_{12}, we investigated their potencies and efficacies (rows 106-110 in tab 'BIG-QualComp' in Spreadsheet S3). We find that unsupported couplings in the Bouvier and Inoue datasets have average pEC_{50} values that are 0.4 and 1.5 log units – 2.5-fold and 32-fold, respectively – lower than the supported couplings. Furthermore, their corresponding average E_{max} values are 18% and 15% lower. This points to the possibility that part of differences across datasets is due to weak couplings which can be difficult to distinguish from basal levels, although the lack of support by GtP may suggest that some couplings are false positives. Both possibilities may be addressed by future datasets or by optimization or development of the profiling platforms. However, until further data becomes available, we recommend a requirement of two independent studies in support of a GPCR-G protein coupling. This is the criterion used herein for the novel couplings and in our online coupling atlas in the GPCRdb database (5). Together these data highlight the importance of large-scale G protein coupling profiling studies using complementary approaches to establish a robust atlas of GPCR coupling selectivity. The Inoue and Bouvier datasets represent the first steps in this direction and there is no doubt that additional studies will be forthcoming.

Coupling map reveals 73% promiscuity within and 49% selectivity across protein families

To enable quantitative correlation of the Bouvier, Inoue and future couplings, we established a data processing protocol giving the highest similarity and correlation of coupling measurements across datasets (Methods). This uses log transformed EC_{50} values and minimum-maximum normalized E_{max} values combined into a unified $\log(E_{max}/EC_{50})$ value and an aggregation of G proteins onto families based on the maximum subtype values (Fig. 4a). Based on this protocol, we created a common coupling map, integrating the Bouvier, Inoue and GtP datasets, which is shown in Fig. 4b and made available in the GPCRdb database hub (5). This unified coupling map establishes that it is possible to obtain comparable quantitative values despite the differences between biosensors and enables quantitative cross-study comparisons herein and in future studies from the field.

To gain insight into their levels of coupling selectivity, we intersected the G protein profiles of all receptors and counted the number of coupling partners for GPCRs and G proteins (panels a, b and c, respectively in Fig. 5). On the G protein family level (topmost in Fig. 5), our analysis spans 90 receptors with data only in GtP and 166 GPCRs with data from the Bouvier and/or Inoue groups – totaling 256 receptors. We require couplings from the Bouvier and Inoue groups to be supported by a second dataset (GtP couplings are already typically supported by several publications). We find that a GPCR couples to on average 1.7 G protein families distributed as 126 single-, 83 double-, 34 triple- and 13 all-family activating receptors (49%, 32%, 13% and 5%, respectively, Fig. 5b). The share of fully selective (single-family activating) receptors differs largely across G protein families spanning from 6% for $G_{12/13}$ to 22% for $G_{q/11}$, 26% for G_s and up to 40% for $G_{i/o}$ (tab 'Fig_5' in Spreadsheet S5). Interestingly, all fully promiscuous receptors are class A GPCRs: adenosine A_1, adrenergic $\alpha_{1A,2A}$ and β_1, bradykinin B_2, cannabinoid 1, cholecystokinin CCK_1, endothelin ET_A, prostanoid FP, GPR4, histamine H_1, lysophospholipid LPA_4 and orexin OX_2 receptors (216 (76%) class A, 11 (61%) class B1 and 5 (23%) class C GPCRs have been profiled, so far). Conversely, a G protein family has supported couplings to on average 112 GPCRs (28% of all receptors) distributed as G_s: 87, $G_{i/o}$ 176, $G_{q/11}$: 134 and $G_{12/13}$: 49 receptors (34%, 69%, 52% and 19%, respectively of all receptors, Fig. 5c). Given that 101 of the GPCR-G protein family pairs tested by the Bouvier and Inoue groups represent novel couplings (above) more couplings are expected to be identified as expanding and confirmatory studies emerge. Hence, whereas the results described here represent the currently known supported couplings, the total 'couplome' will undoubtedly comprise additional yet undetected and unconfirmed couplings, especially among receptors never profiled with a pan-G protein platform.

Within each G protein family (rows 2-5 in Fig. 5), on average 73% of GPCRs promiscuously activate all its members ($G_{q/11}$: 73%, $G_{12/13}$: 66%, $G_{i/o}$: 75% and G_s: 80%). In contrast, activation of only one subtype of a G protein family is only observed for 11 G_s, 4 G_z, 1 G_{14}, 10 G_{15}, 5 G_{12} and 10 G_{13}-coupled receptors (Fig. 4 or Spreadsheet S5). Most other receptors are activated by a subset of G proteins in each family. Strikingly, $P2Y_{1,4}$ (G_{15}), $P2Y_{14}$ (G_z) and GPR55 (G_{13}) are fully selective also when considering G proteins from all families i.e., they only couple to a single of the 16 human G proteins (1.4*SD cut-off applied, 6 G_s-coupling receptors are left out, as G_{olf} has so far only been tested by the Inoue group). However, the three purinergic receptors have additional couplings although not supported by a second dataset ($P2Y_4$ and $P2Y_{14}$) or above the 1.4*SD cut-off ($P2Y_1$ was below) and it is possible that as the characterizations expand even fewer, or no receptors are found to engage only a single G protein.

In all, our meta-analysis of GPCR-G protein selectivity reveal a high coupling diversity spanning signaling via a single effector to physiological ligand-receptor systems able to activate all four G protein pathways. Such selective or combined profiles, in interplay with differential spatiotemporal expression (3), can be critical to achieve a specific physiological effect.

Pharmacological classification of G proteins agrees with phylogeny and reveals inverse G_s-$G_{i/o}$ correlation and G_{15} atypicality

To assess whether the evolutionary classification dividing G proteins into four families also represents their pharmacology, i.e., couplings to receptors, we made pairwise correlations of G proteins using Jaccard indices (% of couplings to the same GPCRs) and average log(E_{max}/EC_{50}) differences for the receptors that couple both G proteins (Fig. 6a-b). By combining these two measures and calculating Pearson standard correlation coefficients (Fig. 6c), we find that the strongest correlations are seen for G proteins belonging to the same G protein family (G_s: G_s-G_{olf}: 0.79, $G_{12/13}$: G_{12}-G_{13}: 0.79, G_{io}: min: G_{i1}-G_z: 0.73 and max: G_{i1}-G_{i2}: 0.99, and $G_{q/11}$: min: G_q-G_{15}: 0.73 and max: G_q-G_{11}: 0.99). Altogether, these findings provided confirmation that the phylogenetic relationships typically used to classify G proteins into families also reflect their pharmacological relationships.

G_{15} shares the largest fraction of receptor couplings with other members of the $G_{q/11}$ family but this fraction is relatively low compared to other G protein families and the log(E_{max}/EC_{50}) differences of G_{15} vs other G_q pairs are higher than observed for any other intra-family comparison (Fig. 6b). This is reflected in G_{15} having the lowest mean Pearson standard correlation coefficient to other members of the $G_{q/11}$ family (0.76 compared to 0.90-0.99 for the other subtypes) but similar to G_s and G_{olf} (0.79) and to G_{12} and G_{13} (0.79). G_{15} also has a larger overlap than all other G proteins in its receptor couplings to another G protein family, the $G_{i/o}$ family (Fig. 6a). These findings confirm that G_{15} is a member of the $G_{q/11}$ family although with a weaker relationship and suggests that this is due to a combination of differential receptor coupling selectivity and a activities.

When aggregating the receptor couplings and log(E_{max}/EC_{50}) values for the four G protein families (Fig. 6d-e), the $G_{i/o}$ and $G_{q/11}$ families most frequently couple to the same receptors but have a relatively high average difference in log(E_{max}/EC_{50}) levels, which could contribute to selectivity. In contrast, the G_s and $G_{i/o}$ families most rarely couple to the same receptors and also have the highest average difference in log(E_{max}/EC_{50}) levels. In Fig. 6c, this is manifested as an inverse correlation between G_s and G_{olf} to members of the $G_{i/o}$ families (average cross-family Pearson standard correlation coefficient of -0.31 and -0.17, respectively). Together, this shows that G protein family pairs have evolved differing levels of receptor co-coupling or selectivity, which they have achieved by weakening or abolishing the binding to the same GPCRs.

Discussion

Given that researchers are now faced with three large G protein coupling datasets varying in coverage and couplings, our first objective was to establish a common normalization protocol making quantified values comparable across data sources and a map of all couplings and their experimental support. Our meta-dataset analysis of receptors included in all three sources, revealed 101 novel couplings supported by both the Bouvier and Inoue datasets demonstrating the value of dedicated large scale profiling studies. Furthermore, it revealed a large number of yet unsupported couplings, especially for the G_s, G_{12} and G_{15} subtypes (Spreadsheet S3) – warranting the generation of additional supporting datasets. The choice of protein systems and biosensors for such future studies should take into account the relative strengths and limitations of the profiling platforms used (3), while future studies could benefit from other recently available biosensors (7, 8, 14) (Gαβγ sensors first described in (15-20)) to widen the approaches used to establish consensus couplings. Moreover, tissue-specific couplings (21), spatiotemporal-dependent interactions and unproductive couplings (22) as opposed to actual down-stream signaling events (23) need to be considered when inferring coupling profiles to physiology and pathophysiology. In all, while our analyses point to a need for further studies to confirm couplings detected by a single biosensor, they establish a sizeable reference dataset of supported quantitative couplings suitable for any study.

Previous analyses of GPCR-G protein selectivity have all been based on a single dataset from the GtP (24) database or the Inoue (2) or Bouvier (3) laboratories, whereas the analysis herein combined these datasets to establish couplings supported by two studies (including GtP annotations of multiple literature reports) for a total of 256 GPCRs. This revealed that very few receptors, 13 (5%) promiscuously couple to all four G protein families whereas the receptor numbers instead increase several-fold along with selectivity: 34 triple-, 83 double- and 126 single-family coupling receptors (Fig. 5b). We found that the highest selectivity exists among G_s- versus $G_{i/o}$-coupled GPCRs which most rarely co-couple and, when they do, exhibit the highest average difference in log(E_{max}/EC_{50}) levels (Fig. 6). Given that G_s stimulates, and $G_{i/o}$ inhibits production of the same cellular second messenger, cAMP the selective activation of different receptors at different times presents a very plausible mechanism to modulate cellular cAMP levels with temporal selectivity. Differential engagement of the $G_{i/o}$ family could also have functional consequences through other pathways, as G_i was recently shown to be required for scaffolding β-arrestin binding and signaling for some receptors (25).

In contrast, the $G_{i/o}$ and $G_{q/11}$ families most frequently couple to the same receptors and they also have the largest number of overall receptor couplings: 176 and 134, respectively (compared to 87 G_s and 49 $G_{12/13}$- coupled GPCRs). The abundant coupling of $G_{i/o}$- and $G_{q/11}$ to many receptors with dual (or more) pathways suggests that these G proteins have more versatile functions. Their high log(E_{max}/EC_{50}) differences (Fig. 6e) that allow for a selective activation, which could shift upon the presentation of alternative endogenous or surrogate ligands with a signaling bias. Furthermore, the finding that on average only 27% of GPCRs activate only a subset of the G proteins in each family (Fig. 5) opens the possibility that signaling bias on the level of specific

subtypes may be much more frequent than currently appreciated. G proteins that belong to the same family can have diverse functional outcomes pertaining to effector engagement selectivity and kinetic profiles (26, 27), suggesting that bias within a G protein family may have physiological and/or therapeutic implications (28). The observation that many GPCRs can couple to more than one family of G proteins but yet can show selectivity between members of a same family, opens important questions concerning the structural determinants of such multifaceted selectivity profiles.

In all, our cross-dataset analysis has established a protocol and reference set aiding GPCR-G protein coupling studies. The selectivity profiles are the most comprehensive to date spanning 256 receptors with a very diverse activation of a single to all G protein pathways, and presented in a dedicated online atlas of G protein couplings integrated in the GPCRdb database hub (5). The analyses and data presented herein will be very valuable to further our understanding of undercharacterised pharmacological phenomena such as constitutive activity (29), pre-coupling of G proteins (16, 30) and ligand-dependent biased G protein signaling (31), and to uncover their underlying determinants. They also present the foundation to integrate more coupling data as future studies expand the characterization of the 'couplome'.

Methods

Study design

The primary objective of this study was to generate a unified map of G protein couplings (Fig. 4) across the three available large datasets from the Bouvier (3) and Inoue (2) groups, and the Guide to Pharmacology database (1), respectively. To do this, we identified the Emax standard deviation cut-off, quantitative normalization protocol and aggregation of G proteins into families giving the best possible agreement between the GPCR-G protein couplings from the Bouvier and Inoue groups. This analysis also involved assessing the agreement of the Bouvier and Inoue group datasets and determining the number of high-confidence novel couplings supported by these two datasets but not reported in the Guide to Pharmacology database. Furthermore, it included a benchmarking of techniques to determine which biosensor produces the most reproducible qualitative coupling determination (coupling vs. non-coupling) for each G protein family. To this end, we compared the three datasets to pinpoint, for each dataset and G protein family, the fraction of couplings supported by another dataset. The scientific analyses based on the map feature the most comprehensive analysis to date of GPCR-G protein selectivity. This was done by intersecting the 11 G proteins tested by both Bouvier and Inoue, within and across their families, with respect to the common and unique receptors that they couple to. Finally, the receptor profiles were also used to classify G proteins and to determine any co-correlation between different G protein subtypes and families to GPCRs.

Coupling datasets

Updated Spreadsheets containing the pEC_{50}, E_{max}, basal signals and standard deviation (SD) values were supplied by Asuka Inoue. Basal signals, spontaneous AP-TGF-α elease (in % of total AP- TGF-α expression) were recalculated from raw data of the previous coupling-profiling campaign (2) and their SD values were computed from independent experiments ($n \geq 3$). This file contains the neurotensin 1 (NTS_1) and Thyrotropin-releasing hormone (TRH_1) receptors not included in the previous publication (2). In the Inoue dataset, the protease-activated receptors PAR3-4 had negative pEC_{50} values (concentration of mU/ml because the ligand, thrombin, was supplied with its enzymatic activity). For the easiness of integration into the coupling map, we added a value of 10 to their pEC50 values. Data qualities (sigmoidal curves) for the individual GPCR-G protein pairs were manually inspected and concentration-response curves that did not converge nor exceed a threshold (typically, 3% AP-TGF-α release) were regarded as no activity.

The Bouvier dataset ($n \geq 3$) contained some datapoints that were included or excluded based on dedicated analyses. Firstly, we excluded ligand-promoted responses of overexpressed receptors that were equivalent to those of endogenously expressed receptors (yellow fill in tab 'B-EmEC' in Spreadsheet S1). Secondly, couplings with only approximate E_{max} and pEC_{50} values because the dose-response curve did not converge were only included if supported, and not contradicted, by the Inoue and/or GtP datasets (orange fill in tab 'B-EmEC' and analyzed separately in tab 'UnconvergedDRV' in Spreadsheet S1). This is because a coupling with an unprecise quantitative value is better than no coupling, especially when making qualitative comparisons (coupling vs. non-coupling). Based on these criteria seven couplings were included: 5-HT_{1D}-G_z, BLT_1-G_{11}, FFA3-G_{13}, GPR_4-Gi_{1-2}, GPR84-G_{oA} and κ-G_{12}, and six couplings were excluded: CCR5-G_{14}, CXCR5-G_q/G_{11}/$_{G14}$, GPR183-G_{13} and κ-G_{13}.

Standard deviation cut-off

We made a special investigation of couplings that have a full dose-response curve but an E_{max} less than 2 standard deviations (SDs) from the basal signal (red fill in tab 'B-EmEC' and analyzed separately in tab tabs 'B-<2SDs' and 'I-<2SDs' in Spreadsheet S1). To achieve the best possible separation of putative false and real but weak couplings, we identified the threshold value (a number of SDs from basal signal) that gives the best agreement between the Bouvier and Inoue dataset among the common receptors and

G proteins tested by both groups. The obtained cut-off, 1.4 SDs was applied as a filter to exclude all Bouvier and Inoue E_{max} and pEC_{50} values for couplings below this cut-off (columns with '>1.4SD' in the heading in tabs 'B-EmEC' and 'I-EmEC' in Spreadsheet S1). As a note, while intra-day measurement error is small for the TGF-α shedding assay (typically, 1-2% AP-TGF-α release), inter-day variability varies widely depending on cell conditions. Since the SD represents inter-day variability, the basal SD cut-off removes more couplings than the SD cut-off of ligand-induced signal or the significance of individual experiments. As a consequence, some of manually annotated couplings in the Inoue dataset are regarded as non-coupling by the basal SD cut-off criteria, including for $P2RY_2$ and $P2RY_6$ that had no couplings above this cut-off.

Generating a subset of comparable GPCR-G protein couplings

To enable qualitative comparison of corresponding datapoints in the datasets from the Bouvier (32) and Inoue (2) groups, we used the subset of 70 GPCRs present in both datasets and belonging to the same class, A (removed only two receptors from class B1). The quantitative comparisons focused on a smaller subset of 51 GPCRs tested with the same ligand and excluded non-coupling GPCR-G protein pairs, as they could not be represented by 0 values (due to e.g., an underrepresentation of couplings in a given datasets and G proteins, see Results). All analyses herein included the 12 G proteins: G_s, G_{i1}, G_{i2}, G_{oA}, G_{oB}, G_z, G_q, G_{11}, G_{14}, G_{15}, G_{12} and G_{13}. G_{olf} and G_{i3} could not be analyzed, as they had not been tested by the Bouvier group (Supplementary Table 1). The Inoue data for the pairs G_{i1-2}, G_{oA-B} and G_q and G_{11}, respectively, were generated with identical chimera inserting the Gα C-terminal hexamer into a G_q backbone (2). Qualitative analyses compared the presence or absence of each GPCR-G protein coupling while the quantitative analyses were limited to common G protein couplings, i.e., data points in which both the Bouvier and Inoue groups generated a pEC_{50} and E_{max} value.

Validation of comparability when E_{max} or reference agonist differ

To get an overview of the distribution of data in the two datasets, we determined the pEC_{50} and E_{max} mean, median, min, max, span and standard deviation values and plotted box and whiskers plots for each G protein (Spreadsheet S1). This showed that the E_{max} values vary much more across the different G proteins than pEC_{50} values. The E_{max} variation is largest in the Bouvier data wherein G_{15} ranges across three orders of magnitude (16 to 1,067). The Bouvier dataset has low means and spans for G_s (42 and 47) and G12 (71 and 101) indicating a narrow assay signal window (low signal-to-noise). We find that minimum-maximum normalization (to 100%) of each G protein across receptors gives a more uniform distribution (SDs for Bouvier: 19-31 and Inoue: 22-26, rightmost plot pair in Spreadsheet S1).

Whereas both studies tested a majority of receptors with their endogenous ligand, surrogate agonists were used for 15 and 4 GPCRs in the data from the Inoue and Bouvier groups, respectively. Although those ligands were selected for their reference character with similar pharmacology to the endogenous ligand, they could introduce differences in a receptor's G protein profile due to ligand-dependent signaling bias. Analysis of GPCR-G protein couplers and non-couplers (tab 'BI' in Spreadsheet S2) shows that on average 74% and 71% agreeing qualitative couplings (i.e., coupling vs. non-coupling GPCR-G protein pairs) for receptors tested with the same and different agonists, respectively. This is a rather small difference providing confirmation that receptors tested with different ligands can be compared on the qualitative coupling/non-coupling level and be included in the comparison of the different G protein coupling datasets and determination of novel G protein couplings (Fig. S1).

Dataset integration into a unified coupling map – Normalization and $\log(E_{max}/EC_{50})$ values

To enable quantitative correlation of the Bouvier and Inoue couplings, we further filtered the 70 common GPCRs to yield 51 common class A GPCRs tested with the same ligand (excluding 29 receptors tested with different ligands and two class B1 GPCRs). To assess the value of normalization, we calculated the average 'similarity' (Bouvier/Inoue ratio) of individual values (tabs ending with '-sim' in Spreadsheet S4) and the 'linear correlation' (r2 value) of each receptor across all G proteins (reported below as averages of individual couplings and G proteins, respectively) (tabs ending with '-r2' in Spreadsheet S4). Linear correlation was only done for receptors with at least three common G proteins/families. Minimum-maximum normalized Emax values were represented as percentage values while decimal values (0-1) were used for the calculation of log(Emax/EC50) values as recommended in (4). For both the Bouvier and Inoue groups, the minimum and maximum represent the signal without (0%) and with (100%) an agonist, respectively, in each experiment replica. The minimum Emax value was therefore set 0 while the maximum was set to the highest value for the given G protein (first, use normalization) or receptor (second, tested but not used normalization).

Minimum-maximum normalization of E_{max} measurements increased the average from 0.22 to 0.61 and the linear correlation r^2 value from 0.28 to 0.30. In contrast, non-normalized pEC_{50} measurements have the most similar values (0.89 compared to 0.66) and an identical linear correlation (0.37). We combined the minimum-maximum normalized E_{max} and non-normalized EC_{50} values into $\log(E_{max}/EC_{50})$ (4). The $\log(E_{max}/EC_{50})$ values have an average similarity that is nearly as high (0.86 compared to 0.89) and an average linear correlation that is better (0.41 compared to 0.37) than for pEC_{50}, the best individual measure. We did not apply a

double normalization, i.e., also across G proteins after across receptors, as this worsens the E_{max}, pEC_{50} and $\log(E_{max}/EC_{50})$ value similarity (from 0.61 to 0.39, 0.89 to 0.42 and 0.86 to 0.47, respectively) and linear correlation of each G protein across receptors (from 0.11 to 0.05, 0.49 to 0.20 and 0.46 to 0.23, respectively.

Dataset integration into a unified coupling map – G protein family aggregation

Given that G proteins belong to families that are functionally grouped by sharing downstream signaling pathways, we next investigated the best member-to-family aggregation scheme – specifically, whether the most comparable G protein family values are obtained if using the maximum value from any single subtype or the mean of all subtype members. We found that aggregation using max rather than mean values gives a better similarity for E_{max} (0.69 vs. 0.65), marginally lower similarities for pEC_{50} (0.89 relative 0.90) and identical similarities for $\log(E_{max}/EC_{50})$ (0.88). However, max performs better overall than mean in the comparisons correlating a receptor across G proteins; E_{max} (0.63 and 0.53), pEC_{50} (0.56 and 0.56) and $\log(E_{max}/EC_{50})$ (0.65 and 0.59) or one G protein across receptors; E_{max} (0.29 and 0.20), pEC_{50} (0.71 and 0.74) and $\log(E_{max}/EC_{50})$ (0.69 and 0.71). Notably, the correlation of each receptor across max-aggregated G protein families compared to non-aggregated subtypes is much stronger when considering any of the three pharmacological parameters: E_{max} (0.63 vs. 0.36), pEC_{50} (0.56 vs. 0.30) and $\log(E_{max}/EC_{50})$ (0.65 vs. 0.41). Altogether, this establishes the highest value (max) as the aggregation that yields the most comparable quantitative value for G protein families.

Statistical analysis

The aggregated sample size is n=3 or higher for all analyzed GPCR-G protein couplings. The specific sample size for each such coupling is described in the original articles reporting these data and referenced in the present manuscript. The Bouvier dataset contained some datapoints that were included or excluded based on dedicated analyses (see 'Coupling datasets' above). For all figures, the values for *N*, *P*, and the specific statistical test is included in the figure legend or associated manuscript text. For fig. 6, pairwise correlation of log(Emax(EC50) values was assessed as a measure of the strength of the linear relationship between two pathways. The pairwise distances between pathways were calculated in python scipy package using the *spatial.distance.pdist* method using the *correlation* metric. Statistical significance was determined by Pearson standard correlation coefficient with a two-tailed p-value as implemented in scipy's *stats.pearsonr* method.

References

1. Armstrong JF, Faccenda E, Harding SD, Pawson AJ, Southan C, Sharman JL, et al. The IUPHAR/BPS Guide to PHARMACOLOGY in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY. Nucleic Acids Res. 2020;48(D1):D1006-D21. 10.1093/nar/gkz951
2. Inoue A, Raimondi F, Kadji FMN, Singh G, Kishi T, Uwamizu A, et al. Illuminating G-Protein-Coupling Selectivity of GPCRs. Cell. 2019;177(7):1933-47 e25. 10.1016/j.cell.2019.04.044
3. Avet C, Mancini A, Breton B, Le Gouill C, Hauser AS, Normand C, et al. Effector membrane translocation biosensors reveal G protein and B-arrestin profiles of 100 therapeutically relevant GPCRs. Accompaying manuscript.
4. Kenakin T. A Scale of Agonism and Allosteric Modulation for Assessment of Selectivity, Bias, and Receptor Mutation. Molecular pharmacology. 2017;92(4):414-24. 10.1124/mol.117.108787
5. Pandy-Szekeres G, Kooistra AJ, Esguerra M, Hauser AS, Caroli J, Munk C, et al. The G protein database, GproteinDb. Submitted.
6. Quoyer J, Janz JM, Luo J, Ren Y, Armando S, Lukashova V, et al. Pepducin targeting the C-X-C chemokine receptor type 4 acts as a biased agonist favoring activation of the inhibitory G protein. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(52):E5088-97. 10.1073/pnas.1312515110
7. Olsen RHJ, DiBerto JF, English JG, Glaudin AM, Krumm BE, Slocum ST, et al. TRUPATH, an open-source biosensor platform for interrogating the GPCR transducerome. Nat Chem Biol. 2020;16(8):841-9. 10.1038/s41589-020-0535-8
8. Maziarz M, Park JC, Leyme A, Marivin A, Garcia-Lopez A, Patel PP, et al. Revealing the Activity of Trimeric G-proteins in Live Cells with a Versatile Biosensor Design. Cell. 2020;182(3):770-85.e16. 10.1016/j.cell.2020.06.020
9. Malfacini D, Patt J, Annala S, Harpsoe K, Eryilmaz F, Reher R, et al. Rational design of a heterotrimeric G protein alpha subunit with artificial inhibitor sensitivity. The Journal of biological chemistry. 2019;294(15):5747-58. 10.1074/jbc.RA118.007250
10. Blank JL, Ross AH, Exton JH. Purification and characterization of two G-proteins that activate the beta 1 isozyme of phosphoinositide-specific phospholipase C. Identification as members of the Gq class. The Journal of biological chemistry. 1991;266(27):18206-16.
11. Tomura H, Itoh H, Sho K, Sato K, Nagao M, Ui M, et al. Betagamma subunits of pertussis toxin-sensitive G proteins mediate A1 adenosine receptor agonist-induced activation of phospholipase C in collaboration with thyrotropin. A novel stimulatory mechanism through the cross-talk of two types of receptors. The Journal of biological chemistry. 1997;272(37):23130-7. 10.1074/jbc.272.37.23130
12. Jiang H, Kuang Y, Wu Y, Smrcka A, Simon MI, Wu D. Pertussis toxin-sensitive activation of phospholipase C by the C5a and fMet-Leu-Phe receptors. The Journal of biological chemistry. 1996;271(23):13430-4. 10.1074/jbc.271.23.13430
13. CAMPS M, HOU C, SIDIROPOULOS D, STOCK JB, JAKOBS KH, GIERSCHIK P. Stimulation of phospholipase C by guanine-nucleotide-binding protein βγ subunits. Eur J Biochem. 1992;206(3):821-31. https://doi.org/10.1111/j.1432-1033.1992.tb16990.x
14. Masuho I, Martemyanov KA, Lambert NA. Monitoring G Protein Activation in Cells with BRET. Methods Mol Biol. 2015;1335:107-13. 10.1007/978-1-4939-2914-6_8

15. Galés C, Rebois RV, Hogue M, Trieu P, Breit A, Hébert TE, et al. Real-time monitoring of receptor and G-protein interactions in living cells. Nat Methods. 2005;2(3):177-84. 10.1038/nmeth743
16. Galés C, Van Durm JJ, Schaak S, Pontier S, Percherancier Y, Audet M, et al. Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes. Nat Struct Mol Biol. 2006;13(9):778-86. 10.1038/nsmb1134
17. Schrage R, Schmitz AL, Gaffal E, Annala S, Kehraus S, Wenzel D, et al. The experimental power of FR900359 to study Gq-regulated biological processes. Nat Commun. 2015;6:10156. 10.1038/ncomms10156
18. Breton B, Sauvageau É, Zhou J, Bonin H, Le Gouill C, Bouvier M. Multiplexing of multicolor bioluminescence resonance energy transfer. Biophys J. 2010;99(12):4037-46. 10.1016/j.bpj.2010.10.025
19. Bünemann M, Frank M, Lohse MJ. Gi protein activation in intact cells involves subunit rearrangement rather than dissociation. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(26):16077-82. 10.1073/pnas.2536719100
20. Janetopoulos C, Jin T, Devreotes P. Receptor-mediated activation of heterotrimeric G-proteins in living cells. Science. 2001;291(5512):2408-11. 10.1126/science.1055835
21. Apostolakou AE, Baltoumas FA, Stravopodis DJ, Iconomidou VA. Extended Human G-Protein Coupled Receptor Network: Cell-Type-Specific Analysis of G-Protein Coupled Receptor Signaling Pathways. J Proteome Res. 2020;19(1):511-24. 10.1021/acs.jproteome.9b00754
22. Qin K, Dong C, Wu G, Lambert NA. Inactive-state preassembly of G(q)-coupled receptors and G(q) heterotrimers. Nat Chem Biol. 2011;7(10):740-7. 10.1038/nchembio.642
23. Okashah N, Wright SC, Kawakami K, Mathiasen S, Zhou J, Lu S, et al. Agonist-induced formation of unproductive receptor-G₁₂ complexes. Proceedings of the National Academy of Sciences. 2020;117(35):21723-30. 10.1073/pnas.2003787117
24. Flock T, Hauser AS, Lund N, Gloriam DE, Balaji S, Babu MM. Selectivity determinants of GPCR-G-protein binding. Nature. 2017;545(7654):317-22. 10.1038/nature22070
25. Smith JS, Pack TF, Inoue A, Lee C, Zheng K, Choi I, et al. Noncanonical scaffolding of G_{αi} and β-arrestin by G protein–coupled receptors. Science. 2021:eaay1833. 10.1126/science.aay1833
26. Jiang M, Bajpayee NS. Molecular mechanisms of go signaling. Neurosignals. 2009;17(1):23-41. 10.1159/000186688
27. Ho MKC, Wong YH. Gz signaling: emerging divergence from Gi signaling. Oncogene. 2001;20:1615-25.
28. Anderson A, Masuho I, Marron Fernandez de Velasco E, Nakano A, Birnbaumer L, Martemyanov KA, et al. GPCR-dependent biasing of GIRK channel signaling dynamics by RGS6 in mouse sinoatrial nodal cells. Proc Natl Acad Sci U S A. 2020;117(25):14522-31. 10.1073/pnas.2001270117
29. Berg KA, Clarke WP. Making Sense of Pharmacology: Inverse Agonism and Functional Selectivity. Int J Neuropsychopharmacol. 2018;21(10):962-77. 10.1093/ijnp/pyy071
30. Civciristov S, Ellisdon AM, Suderman R, Pon CK, Evans BA, Kleifeld O, et al. Preassembled GPCR signaling complexes mediate distinct cellular responses to ultralow ligand concentrations. Science signaling. 2018;11(551). 10.1126/scisignal.aan1188
31. Kenakin T. Biased Receptor Signaling in Drug Discovery. Pharmacol Rev. 2019;71(2):267-315. 10.1124/pr.118.016790
32. Avet C, Mancini A, Breton B, Le Gouill C, Hauser AS, Normand C, et al. Selectivity Landscape of 100 Therapeutically Relevant GPCR Profiled by an Effector Translocation-Based BRET Platform. bioRxiv. 2020:2020.04.20.052027. 10.1101/2020.04.20.052027
33. Kooistra AJ, Mordalski S, Pándy-Szekeres G, Esguerra M, Mamyrbekov A, Munk C, et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. Nucleic Acids Res. 2021;49(D1):D335-D43. 10.1093/nar/gkaa1080

Acknowledgments

Nevin Lambert is acknowledged for fruitful discussions. We thank Ayumi Inoue (Tohoku University) for reanalyzing the raw data of the TGF-α shedding assay to obtain SD values for the individual GPCR-G protein pairs.

Additional information

Funding

Canada Research Chair in Signal Transduction and Molecular Pharmacology (MB); Canadian Institutes of Health Research grant FDN-148431 (MB); Lundbeck Foundation grants R218-2016-1266 and R313-2019-526 (DEG); Novo Nordisk Foundation grant NNF18OC0031226 (DEG); BINDS program grant JP20am0101095 (AI); LEAP program grant JP20gm0010004 (AI); Japan Agency for Medical Research and Development (AMED) (AI); Takeda Science Foundation (AI); Uehara Memorial Foundation (AI);

Author contributions

Conceptualization: DEG; Methodology: DEG; Validation: ASH and DEG; Formal Analysis: AI: CA: ASH and DEG; Investigation: ASH: CA: and DEG; Resources: CG; Data Curation: AM: CA: CN: DEG; Writing – Original Draft: DEG and MB; Writing – Review & Editing: AI: ASH: CA: DEG and MB; Visualization: ASH and DEG; Supervision: DEG: and MB; Project Administration: DEG; Funding Acquisition: AI: DEG and MB;

Data and materials availability

All data are available in GPCRdb (https://review.gpcrdb.org), GitHub (https://github.com/protwis/gpcrdb_data) and Spreadsheets S1-5. No code was developed for this manuscript which instead used the existing GPCRdb (33) resources, including a new G protein resource (5). All open-source code can be obtained from GitHub (https://github.com/protwis/protwis) under the permissive Apache 2.0 License (https://www.apache.org/licenses/LICENSE-2.0).

Competing interests

MB is the president of Domain Therapeutics scientific advisory board. AM and CN were employees of Domain Therapeutics North America during part or all of this research.

Figures and Tables

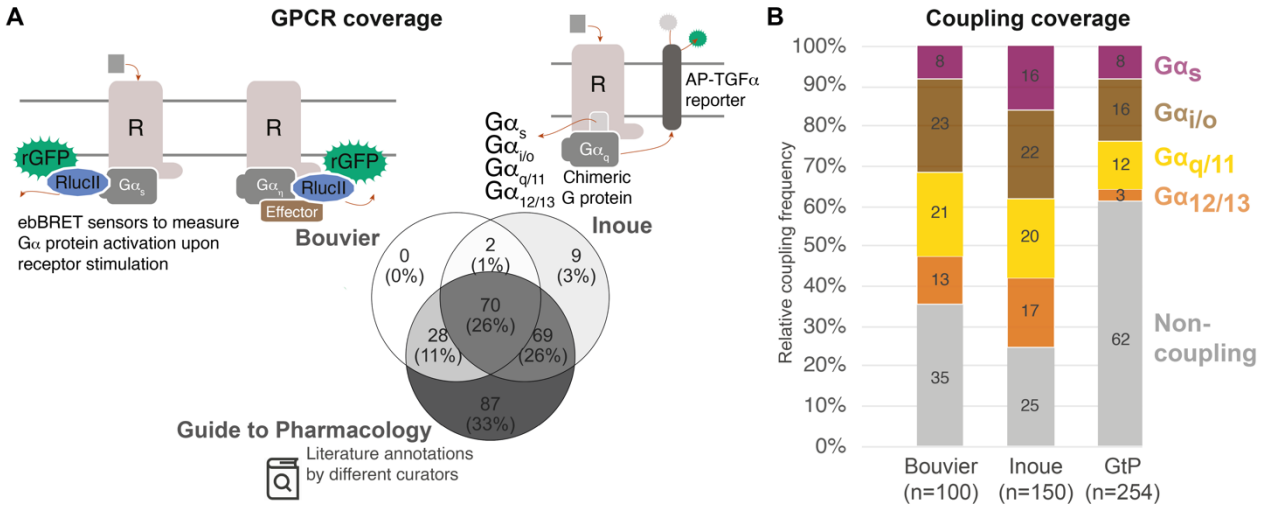


Fig. 1. Bouvier lab, Inoue group and Guide to Pharmacology coupling dataset coverages. Comparison of the G protein coupling profiles generated by the Bouvier (100 GPCRs) and Inoue groups (150 GPCRs) and from the Guide to Pharmacology database (GtP, 254 GPCRs). (**A**) Overlap in receptor coverage between the three coupling datasets. (**B**) Relative family distributions of G protein couplings across datasets.

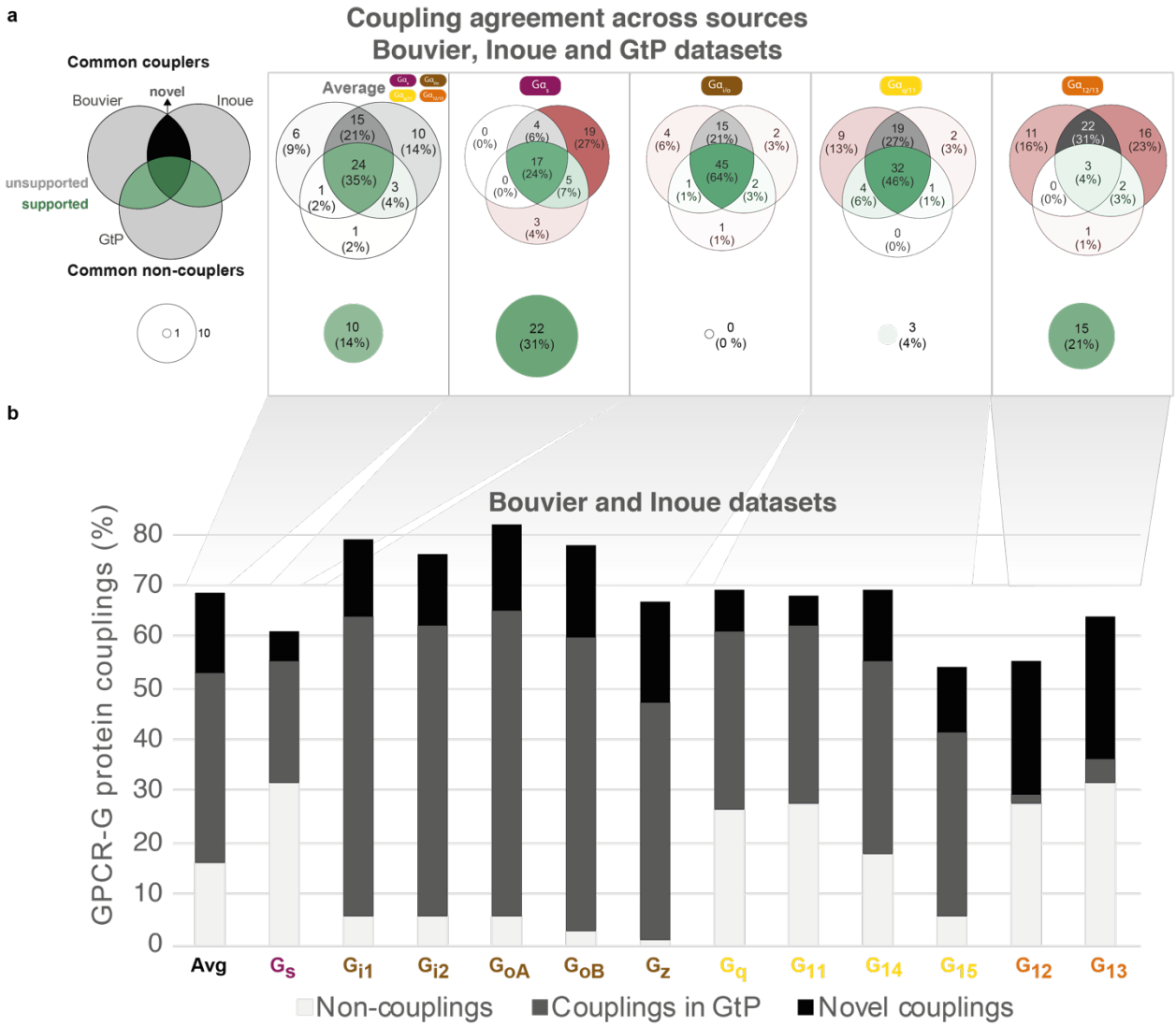


Fig. 2. Bouvier and Inoue group agree on 69% of couplings and 15% are novel. Comparison of the G protein coupling profiles of 70 GPCRs present in all of the Bouvier, Inoue and GtP datasets. G_{i3} and G_{olf} (chimera) are not included herein as they were only tested by the Inoue group. Inoue used the same chimera to represent the pairs G_{oA}-G_{oB}, G_{i1}-G_{i2} and G_q-G_{11}. (Table S1). (**A**) Venn diagrams of the common 'supported' and unique 'unsupported' couplings and bubble diagram of non-couplings. (**B**) Common GPCR-G protein couplings in the Bouvier and Inoue datasets. Black bars represent couplings not in GtP and therefore considered novel (novel couplings are listed in Fig. S1). Frequencies are percent couplings among all receptors common to Bouvier and Inoue (the 100% includes non-coupling receptors).

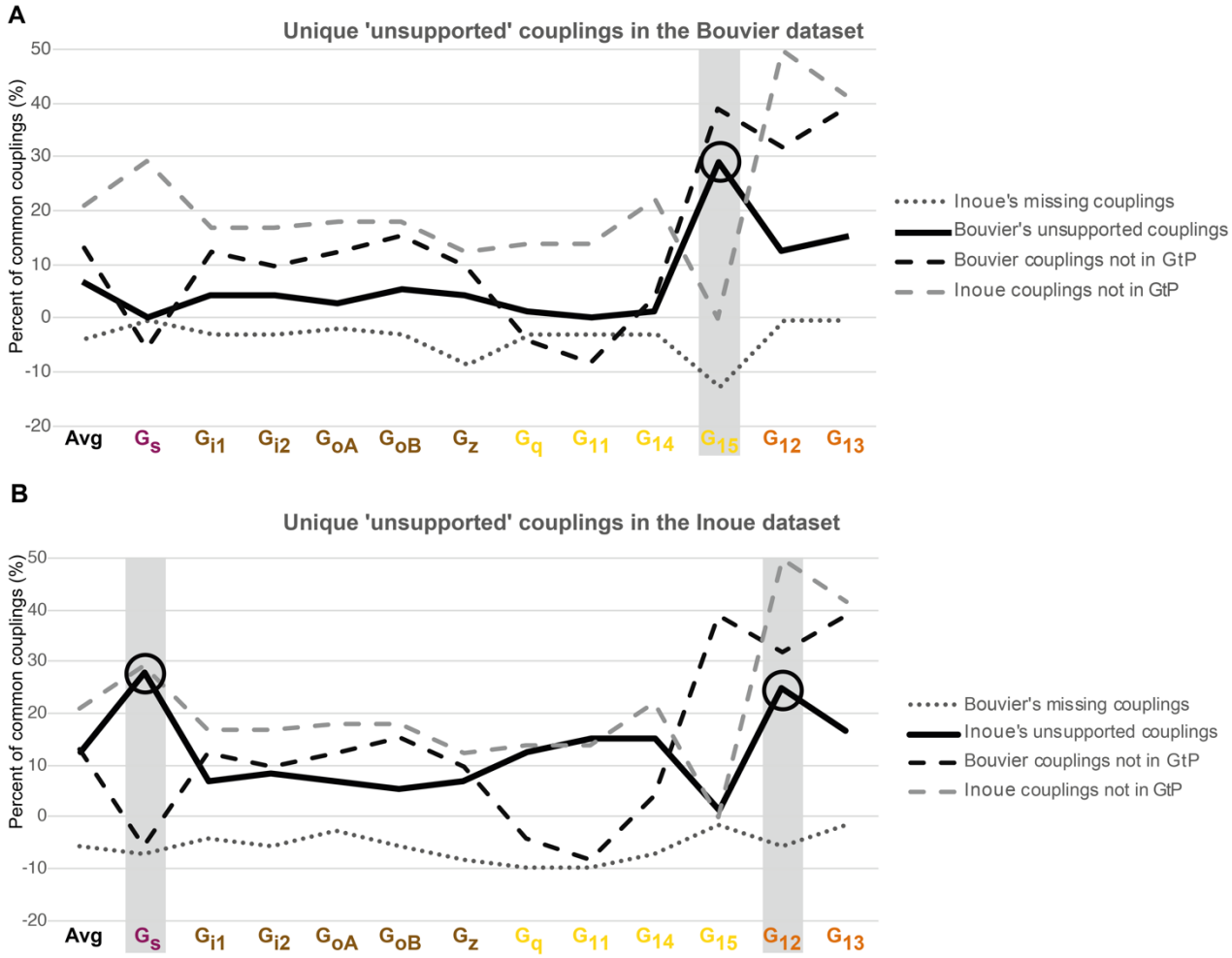


Fig. 3. Unique 'unsupported' couplings in the Bouvier and Inoue datasets. Rationalization of unique 'unsupported' couplings in the (**A**) Bouvier and (**B**) Inoue datasets not observed in the other datasets or in GtP. Frequencies in the graphs are percent couplings among all 70 receptors common to Bouvier, Inoue and GtP (the 100% includes non-coupling receptors). Missing couplings are those present in the two other datasets and are shown as negative values. (**A**) In the Bouvier dataset, G_{15} has large fractions of couplings that are unique (29%) and not available in GtP (39%). In contrast in the Inoue dataset, G_{15} has no new (0%) but several missing (13%) couplings compared to GtP. This indicates that G_{15} couplings are underrepresented in literature and in the Inoue data. (**B**) In the Inoue dataset, G_s and G_{12} have large fractions of couplings that are unique (28% and 25%) and not available in GtP (29% and 50%). In contrast, the Bouvier dataset is missing 7% of G_s and 6% of G_{12} couplings in GtP. The unique Inoue G_s and G_{12} couplings await support from future studies.

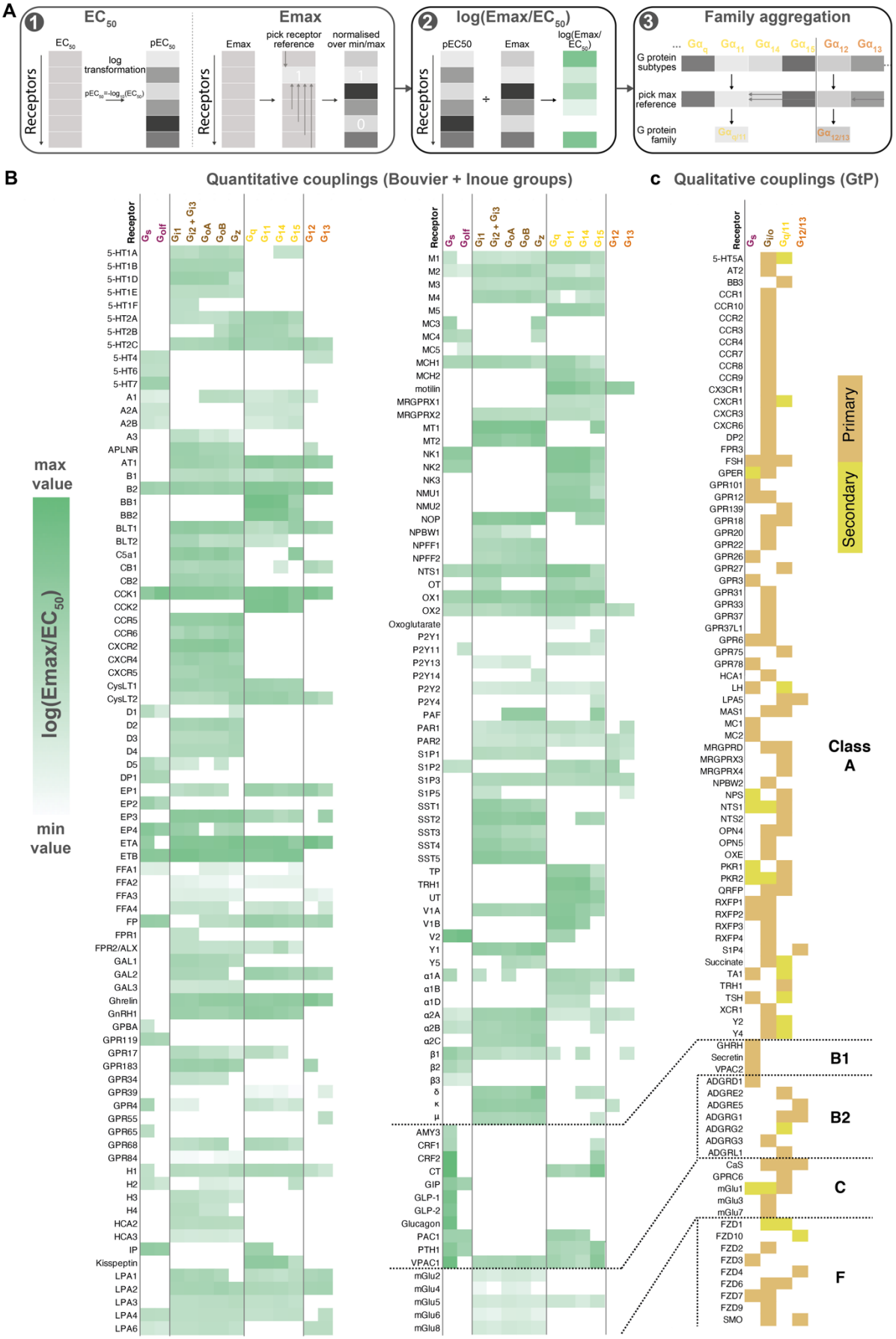
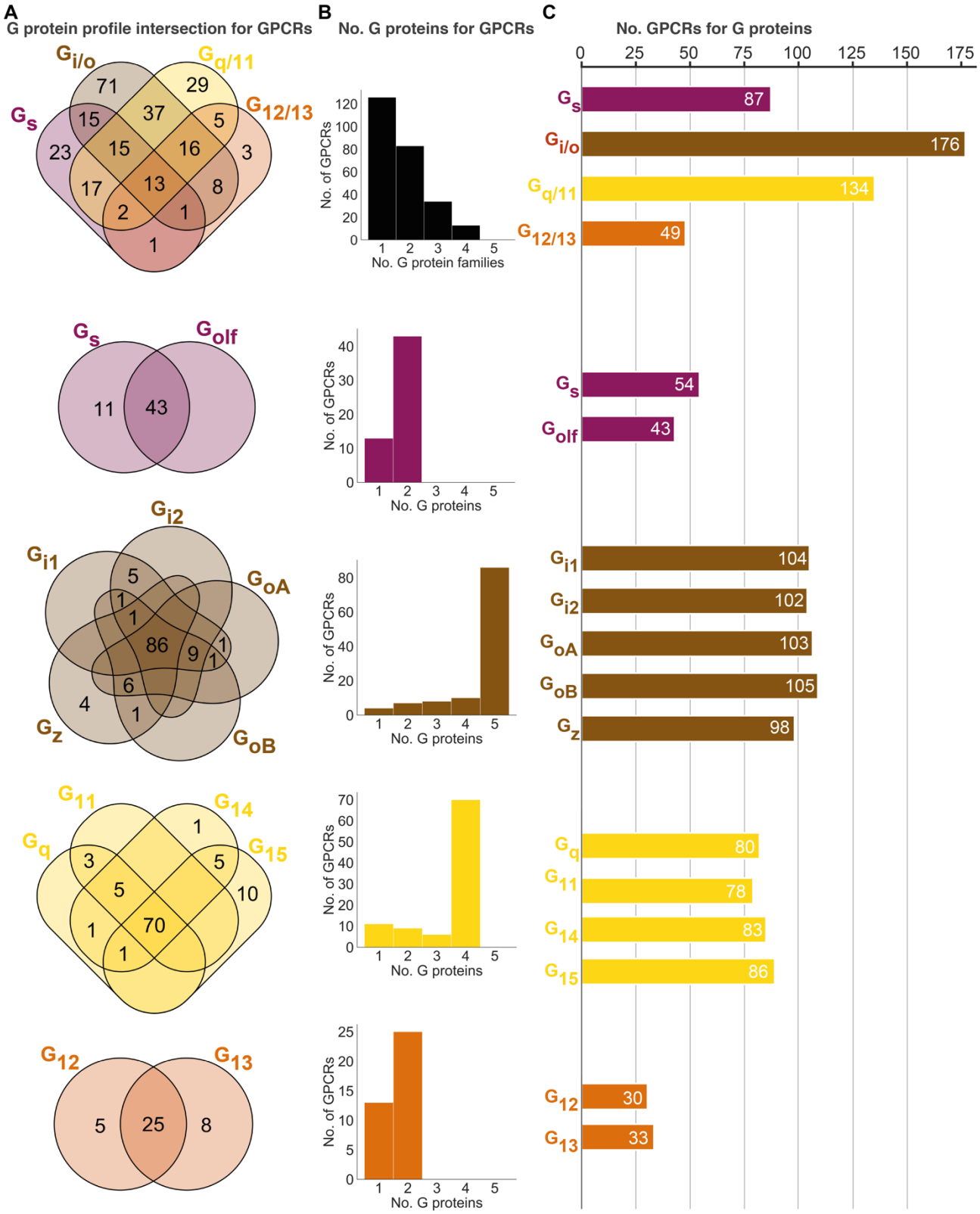

A
1 EC50
Emax
Receptors
log transformation
pEC50=-log10(EC50)
pick receptor reference
normalised over min/max
2 log(Emax/EC50)
pEC50
log(Emax/EC50)
3 Family aggregation
G protein subtypes
pick max reference
G protein family
B
Quantitative couplings (Bouvier + Inoue groups)
C Qualitative couplings (GtP)
Receptor
max value
min value
log(Emax/EC50)
Primary
Secondary
Class A
B1
B2
C
F

Fig. 4. Map of GPCR-G protein couplings supported by at least two studies. (**A**) Normalization approach. Emax minimum-maximum normalization, EC50 log transformation and use of log(Emax/EC50) as a combined measure with member-to-family aggregation by maximum G protein value. (**B**) Heatmap representation of log(Emax/EC50) values for 166 GPCRs tested by the Bouvier and/or Inoue labs (for couplings with dual data sources a mean is used). For GtP, a G protein subtype is considered supported if the family has a known coupling. Gi2 and Gi3 are represented by the same / identical chimeric G protein in Inoue's dataset (Table S1). (**C**) Heatmap representation of primary and secondary transducers for 90 GPCRs which couplings are only covered by the GtP database. (**B-C**) Empty cells (white) indicate no coupling. All source values are available in tab 'Fig_4' in Spreadsheet S5.


A
G protein profile intersection for GPCRs
Gi/o
Gq/11
Gs
G12/13
71
29
15
37
5
23
15
16
3
17
13
8
2
1
1
Gs
Golf
11
43
Gi2
Gi1
GoA
5
1
1
86
9
1
1
6
4
1
Gz
GoB
G11
G14
Gq
G15
1
3
5
5
10
1
70
1
G12
G13
5
25
8
B
No. G proteins for GPCRs
No. of GPCRs
No. G protein families
No. G proteins
C
No. GPCRs for G proteins
0
25
50
75
100
125
150
175
Gs
87
Gi/o
176
Gq/11
134
G12/13
49
Gs
54
Golf
43
Gi1
104
Gi2
102
GoA
103
GoB
105
Gz
98
Gq
80
G11
78
G14
83
G15
86
G12
30
G13
33

08

Fig. 5. Unifying all available data sources reveals a large diversity in GPCR coupling selectivity. (**A**) GPCR-G protein selectivity on the G protein family (top) and subtype levels: Venn diagrams showing the numbers of shared and unique receptors. 0 values (no receptors) are omitted for clarity. (**B**) Receptor coupling promiscuity: Number of receptors that couple to 1 to 4 G protein families (top) and 1 to 5 subtypes. (**C**) G protein coupling promiscuity: Number of receptors that couple to each G protein family (top) or subtype. (**A-C**) Panels **B-C** are based on the couplings from the Bouvier and Inoue groups that are also supported by a second dataset (Fig. 4b) and panel **A** additionally includes GtP couplings (Fig. 4c). This analysis of the G_s family leaves out 11 receptors tested for coupling to G_s but not to G_{olf}, and G_{olf} couplings are only counted if there is a supported G_s coupling. All source data are available in tab 'Fig_5' in Spreadsheet S5.

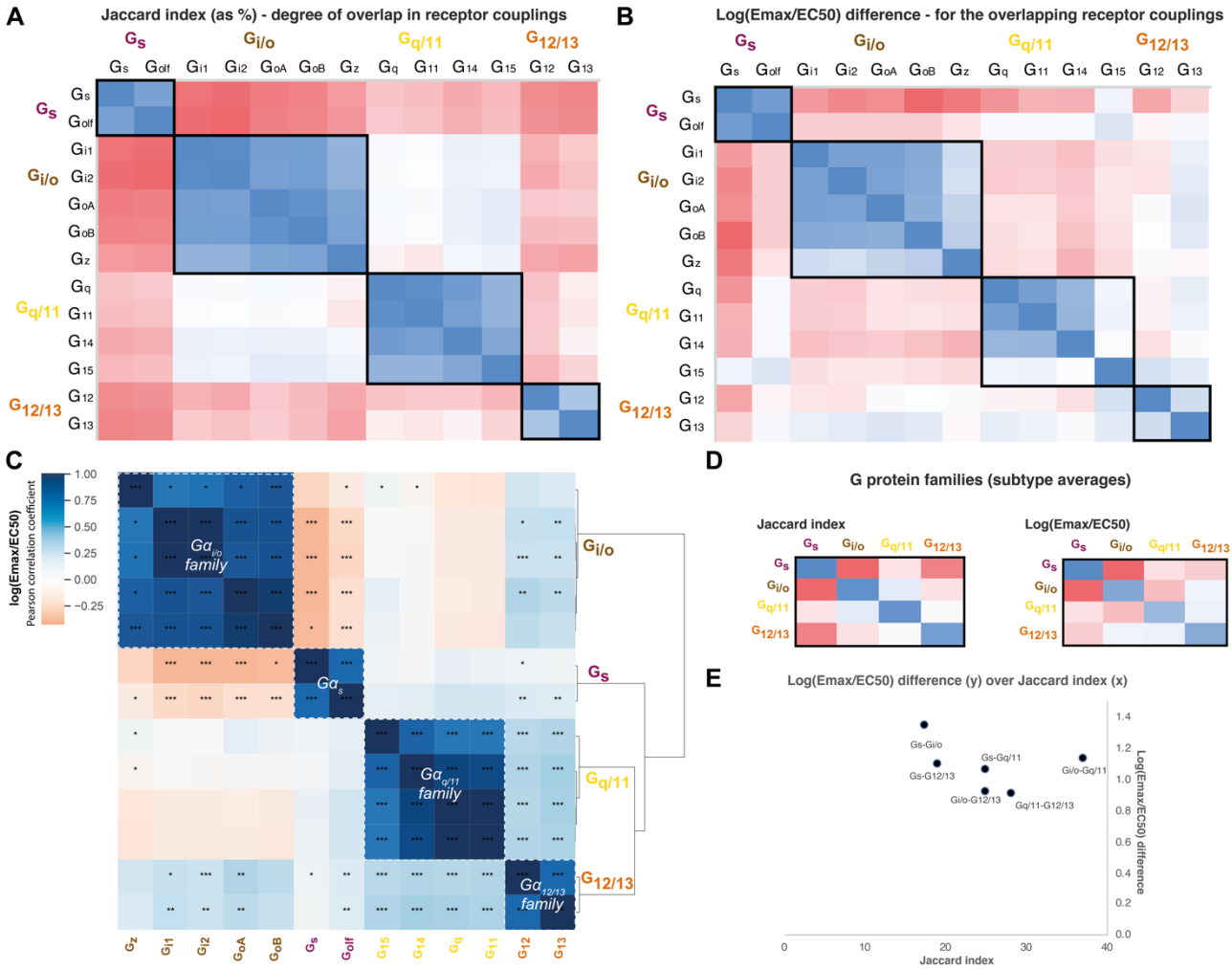


Fig. 6. Pharmacological correlation of G proteins based on their receptor coupling profiles. (**A**) G protein pair Jaccard indices (% of couplings to the same GPCRs). (**B**) G protein pair average log(E_{max}/EC_{50}) differences for the receptors that couple both G proteins. The values are averages of the log(E_{max}/EC_{50}) averages of Bouvier and Inoue, except where data is only available in one dataset i.e., Bouvier only: G_{i1}-G_{i2}, G_{oA}-G_{oB}, G_q-G_{11}, and Inoue only: G_{olf}-all G proteins. (**C**) Pairwise correlation of log(E_{max}(EC_{50}) values (no coupling set to 0) by Pearson standard correlation coefficient, which gives a measure of the strength of the linear relationship between two pathways. Statistically significant pairwise correlations are indicated in cells by *$p \leq 0.05$; **$p \leq 0.005$ and ***$p \leq 0.0005$. (**D**) G protein family pair Jaccard index and average log(E_{max}/EC_{50}) difference calculated as the mean of subtypes in **A**. (**E**) Scatterplot of G protein family Jaccard index (x-axis) and average log(E_{max}/EC_{50}) difference (y-axis). (**A-B**) For G proteins (all but G_{olf}) and receptors tested by both the Bouvier and Inoue groups, an average of averages from the two groups is used. The G_{oA} and G_{oB} couplings from Bouvier were compared to the G_o values from Inoue which does not distinguish isoforms. (**A-C**) Boxes group members of each G protein family. (**A-E**) All G protein couplings are for class A GPCRs and supported by two datasets (Bouvier or Inoue group or GtP), and their source values are available in tab 'Fig_6' in Spreadsheet S5.