The landscape of cancer rewired GPCR signaling axes.

Chakit Arora[1], Marin Matic[1], Natalia De Oliveira Rosa[1], Lorenzo Nemati[1], Lauren Clubb[2], Giorgos Kargas[1], Ranka Vukotic[3], Luana Licata[4], Guanming Wu[5], J. Silvio Gutkind[2] and Francesco Raimondi[1+]

1 Laboratorio di Biologia Bio@SNS, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126, Pisa, Italy; 2 Department of Pharmacology and Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093, USA; 3 Azienda Ospedaliero-Universitaria Pisana, Via Roma, 67, 56126 Pisa;4 Department of Biology, University of Rome 'Tor Vergata', Rome 00133, Italy; 5 Division of Bioinformatics and Computational Biology, Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

+ To whom correspondence should be addressed: francesco.raimondi@sns.it

Abstract

We explored the dysregulation of GPCR ligand signaling systems in cancer transcriptomics datasets. We derived a network of interacting ligands and biosynthetic enzymes from public databases, that we combined with cognate GPCRs and downstream effectors to quantify GPCR signaling pathways. We found multiple GPCRs differentially regulated together with their ligands across cancers, suggesting a widespread perturbation of these signaling axes in specific cancer molecular subtypes. We showed that biosynthetic pathway enrichment from enzyme expression recapitulates pathway activity signatures from metabolomics datasets, therefore providing valuable surrogate information for receptor-organic ligand systems. The expression of several GPCRs signaling components is significantly associated with either lower or higher patient survival in a cancer subtype specific fashion. In particular, the expression of both receptor-ligand (or biosynthetic enzyme) interaction partners improves the stratification of patients based on survival, suggesting a potential synergistic role for the activation of certain receptor axes in modulating cancer phenotypes. Strikingly, we have found that receptors from these actionable axes are the targets of several drugs displaying anti-growth effects in large, drug repurposing screens in cancer cell lines. This study provides a

comprehensive map to exploit GPCR signaling axes as actionable targets for personalized cancer treatments. We have made the results generated in this study freely available for further exploration to the community through a webapp (gpcrcanceraxes.bioinfolab.sns.it).

Introduction

G-protein coupled receptors (GPCRs) are the largest family of transmembrane proteins which transduce a wide number of physico-chemical stimuli into the cell by coupling to intracellular heterotrimeric G-proteins, thereby controlling multiple downstream signaling pathways and transcriptional programs[1–3]. The GPCRs repertoire expanded and diversified to meet the increasing need for communication of the cell in evolving multicellular organisms[4,5]. Systematic efforts are being undertaken to illuminate the complex circuitry of GPCR signaling at both the intracellular [5–10] as well as the extracellular ligand level[11].

Cancer genomics provides the possibility to systematically shed light on cancer somatic alterations affecting GPCRs as well as interacting partners [12–16]. The GPCR-based circuitry, which evolved to support multicellular organization and higher-level functions for eukaryotic cells, is frequently disarranged in cancer. Indeed, A landscape of GPCR-centric autocrine, paracrine, juxtacrine or endocrine signaling networks is emerging as a critical component for the interaction of tumor cells with their tumor micro-environment (TME), ultimately leading to cancer progression and therapy resistance[14,15]. For instance, chemokine receptors are best characterized for their role in the recruitment of immune cells and, depending on the specific cellular context, may results in pro- or anti-tumor effects[14]. Several GPCRs and associated components mediating inflammation have also been connected to immune evasion, such as prostaglandins (in particular *PGE2*), their biosynthesizing enzymes (*PTGS1* and *PTGS2*) and cognate prostanoid receptors. Ectonucleotidases (such as *NT5E* or *ENTPD1*) and Adenosine receptors are also well known to induce an immunosuppressive tumor microenvironment [14]. Recent studies have also shown the emergent role of innervation in promoting tumor growth. Notably, it has emerged a cross-talk between the neurotrophin pathway, particularly the *NGF-NTRK1* axes, and the GPCR ones mediated by either (nora)adrenaline-*ADRB2* in pancreatic cancer[17], and acetylcholine-*CHRM3* in gastric cancer[18].

Single cell and spatial sequencing technologies are allowing us to study mechanisms of cell-cell communications at an unprecedented resolution[19]. In the most comprehensive resource to date used to infer ligand-receptor pairs mediating cell-cell communication[20], GPCRs-ligand interactions are surprisingly depleted, representing the 5% of the total, despite being the most

abundant class of transmembrane proteins. This is likely due to many GPCRs binding to small organic molecules, whose contribution is therefore underestimated in transcriptomics datasets.

In the present study, we considered an extended network of signaling axes formed by GPCRs with interacting ligands, as well as biosynthetic enzymes, to systematically explore their aberrant regulation in cancer transcriptomics datasets and prognostic value for personalized medicine treatments.

Results

GPCRs extracellular signaling system

We generated a comprehensive catalog of upstream and downstream molecular interactions involving GPCRs (Figure1a; see Methods). Briefly, we retrieved known GPCR ligands from IUPHAR/Guide to Pharmacology database (GtoPDB), which we mapped to individual ligand-synthesizing enzymes and known biosynthetic pathways, as well as to pathways mediating signal transduction (Methods; Table S1,2). This yielded a total of 236 ligands, 309 biosynthetic pathways and 128 enzymes associated instances (Figure 1b). We found that (86%) of ligands and (97.65%) of enzymes were already annotated within biological pathways knowledgebases (i.e. Reactome; Figure 1b). A total of 211 (53%) non-olfactory receptors have known ligands in GtoPdb, 95 of which exclusively bind organic ligands, 90 peptide and 26 both kinds (Figure 1b). While most of peptide ligands are annotated to participate in signaling pathways, the organic ones are also extensively annotated in metabolic pathways (Figure 1b; Table S1).

We then computed genome-wide differential expression (DE) by contrasting cancer samples from the TCGA to those from corresponding healthy tissues from GTEx to derive expression signatures of genes and pathways mediating GPCR signaling. We considered a total of 15121 samples, 10534 from TCGA and 4590 from GTEX, encompassing 16 tissues, as well as 76 molecular subtypes compiled from TCGA (see Methods; Table S3).

We found a widespread, cancer-specific regulation of GPCR-related genes with respect to healthy tissues. We found on average 183 and 313 significantly DE genes (FDR < 0.01) across cancers, respectively considering only GPCRs or in combination with ligands and biosynthetic enzymes (Figure 1c). The amount of significantly DE genes varies across tissues, with Testis, Skin and Ovary being the most affected, and Bladder the least (Figure 1c). Certain cancer tissues are characterized by a prevalence of either significantly over-expressed (e.g.

Pancreas) or downregulated (e.g. Brain) GPCRs (Figure S1a). We found no significant trend when grouping DE receptors based on their G-protein coupling specificity (Figure S1a). On the other hand, DE profiling of heterotrimeric G-protein alpha subunits showed that *GNAS* is the single alpha subunit most widely up-regulated, being significantly over-expressed in 94% of cancer tissues (Figure S1b,c), confirming previous analyses[14,16].

We found a widespread DE regulation (FDR < 0.01) of GPCR pathways across cancers (Figure 1d,e). In particular, we found that 40% of cancer tissues are either more enriched in biosynthetic or signal transduction processes, while in 20% of cancer tissues both pathway domains are equally affected (Figure 1d). Several pathways are found to be widespread affected (e.g. up-regulation of Chemokine receptors), while others are exclusively regulated in certain tumor types (Figure 1e). In general, every cancer tissue is characterized by a highly specific GPCR pathway signature (Figure 1e).

The landscape of peptide ligand-GPCRs dysregulation in cancer

We focused our analysis on the extracellular components of the GPCR signaling machinery, i.e. ligands and related metabolizing enzymes, to understand their contribution to the overall dysregulation of GPCR signaling in cancer.

We found an average 66 significantly DE regulated peptide ligand precursors in 16 cancer tissue, and an average 59 instances in 76 molecular subtypes.

We assessed the degree of correlation of expression change between receptors and interacting ligands and found a widespread co-regulation of receptor-ligand precursor axes in multiple cancer tissues or subtypes (Figure 2a, Figure S2). By focusing on the receptor-ligand pairs more recurrently co-regulated, we identified two clusters, showing respectively a prevalence of co-up (i.e. Cluster 1) and co-down (i.e. Cluster 2) axes across subtypes (Figure 2a). The vast majority of most frequently regulated ligands are agonists, with the exception of *CCL4*, *CCL7* and *CXCL8,10,11*, which depending on the bound receptor might act as either agonists or antagonists (Figure 2a). The majority of axes in Cluster 1 involve chemokine receptors (e.g. *CCR1, CCR3, CCR4, CCR5, CCR6, CCR8, CXCR3, CXCR6, XCR1*) as well as other receptors such as *APLNR*, *CALCR*, *GRPR, MCHR1, GPR37, CCKBR, NPBWR1, KISS1R* and cognate ligands. The main transduction mechanism in this group is through $G_{i/o}$ proteins, as well as to a lesser extent via $G_{q/11}$ and $G_{12/13}$ (Figure 2a).

Cluster 2 contains instead receptor-ligand axes that are more frequently co-down regulated in cancer subtypes, although several of these are co-up regulated in specific subtypes (e.g. *CXCR2, EDNRB, FFPR2*, or *LGR6, GALR2* and *UTS2R;* Figure 2a).

The axes in Cluster 2 involve receptors that are members of various families, including: VIP and PACAP, chemokine, Tachykinin, Neuropeptide Y, Class A orphans, Glucagon receptor, Vasopressin and Oxytocin, Formylpeptide, Histamine, Somatostatin, Endothelin, Gonadotrophin-releasing hormone, Melanocortin, Glycoprotein hormone, Cholecystokinin, Galanin and Urotensin. Receptors in this group have a more heterogeneous/promiscuous coupling profile to $G_{q/11}$, $G_{i/o}$ and G_s, but very little to $G_{12/13}$ (Figure 2a).

In general, every cancer subtype is characterized by a highly specific receptor-ligand co-expression signature. However, certain tissues display more similar subtype co-expression patterns (e.g. liver, breast, ovary, thyroid, stomach, lung and skin). Others show higher heterogeneity (e.g. kidney, colon, bladder, esophagus, prostate, brain) (Figure 2a).
Co-regulation patterns of a given receptor can vary, depending on the tissue as well on the number and type of bound ligands (Figure 2b, Figure S2). Certain receptors display consistent co-regulation patterns across cancer tissu0es even if bound to multiple ligands (e.g. *CXCR3, CCR3*, *CCR5*; Figure 2b,2c). Others show more variable co-regulation patterns which vary in a tissue and ligand specific fashion (e.g. *CALCRL* or *RXFP2*; Figure 2b,c).

Biosynthetic pathway enzymes inform on the levels of small molecule ligands

We reasoned that the expression levels of the enzymes synthesizing organic GPCR ligands could be informative of their levels and be used as proxy to model ligand abundances. To this end, we assessed enzyme and biosynthetic pathway activities from differential expression signatures (see Methods).
We found on average 54 and 8 significantly regulated enzymes and biosynthetic pathways, respectively, across cancer, suggesting a widespread dysregulation of GPCR ligand biosynthetic pathways across cancers and specific to distinct subtypes (Figure 3a). Clustering of biosynthetic pathways based on enrichment scores revealed two main clusters: one characterized by processes, such as “Selenocysteine synthesis” or “Interconversion of nucleotide di- or tri- triphosphates”, which are overall up-regulated in the majority of subtypes, while the other is characterized by processes recurrently down-regulated, such as “Phase I - Functionalization of compounds” or “Arachidonic acid metabolism” (Figure 3a).

We compared pathway enrichment scores of DE transcripts with differential abundance scores of DE metabolites from a pancancer metabolomic profiling[21], considering a set of curated biosynthetic pathways (e.g. Reactome; see Methods;Table S4). We found a total of 33 significant pathways (FDR < 0.01) in at least one of the 5 cancer types considered and which

showed concordant pathway enrichment when considering both differentially expressed transcripts and metabolites (Figure 3b). The cancer types showing the highest number of concordant signatures is pancreatic cancer (PAAD). Certain pathways show concordant up-regulation in multiple cancer types, including: “activation of gene expression by SREBF (SREBP)”, “Purine catabolism”, “Pyruvate metabolism”, each found up-regulated in three distinct cancer types (Figure 3b). For example, “Purine catabolism” is always up-regulated when considering both transcriptome and metabolome in Breast, Pancreas and Prostate cancers (Figure 3c). Inspection of differentially expressed genes and metabolites on a functional interaction network shows that over-activation of the pathway is contributed by over-expression of both transcripts and metabolites components (Figure 3e). However, only a few enzymes or metabolite in this pathway are invariably over-expressed across cancers (e.g. *NUDT1*, *LEF1*, 9H-Xanthine; Figure 3d, Figure S3). Other processes show concordant enrichment with a cancer dependent directionality: e.g. nicotinate metabolism, which is up-regulated in pancreas and down-regulated in prostate, or creatine metabolism, down-regulated in pancreas and up-regulated in breast (Figure 3b). Several other pathways are exclusively enriched in specific cancer tissues: for example, “Glucagon-like Peptide-1 (GLP1) regulates insulin secretion”, “Synthesis of Prostaglandins (PG) and Thromboxanes (TX)” and “Acetylcholine regulates insulin secretion” are hallmarks of pancreatic cancer and might contribute to reinforce the known role of the Gs - PKA signaling axis in this cancer type [22,23].

Expression of GPCRs signaling axes components correlates with patient survival

We systematically assessed correlations between over-expression of GPCR signaling components and patient survival.

We first considered the capability of the expression of individual components (e.g. receptors, transducers, ligand or biosynthetic enzymes) to discriminate patients based on their survival. We found a total of 298 unique significant receptor instances (LogRank P-value < 0.01) in 22 cancer tissues, 108 ligands in 23 cancers, 99 enzymes in 17 cancers (Figure 4a; Table S5-7). Some cancer tissues display overall more components whose expression is associated to either higher (e.g. pancreas, skin or head and neck) or lower (e.g. breast, lining of body cavities, white blood cell) survival (Figure 4a). Overall, GPCR expression is associated with poorer survival in the 59% of cases across cancers. *CELSR3* and *LTB4R* are significantly associated with patients’ survival in five cancers each, with a prevalence for lower survival (Figure 4b, Table S5). Some receptors show consistent associations across cancers: for example, *ADORA2A* is invariably associated to higher survival across four distinct cancer tissues (including pancreas, breast, skin and head and neck; Figure 4b), in agreement with their known role in regulating CD8+ T cell activity in the tumor microenvironment[24]. Certain receptors are instead invariably associated with poorer survival, such as *FFAR3*. We also

found a large prevalence of associations between GPCR peptide ligand precursor expression and lower survival of patients, with 67.2% of the significant instances across cancers. Among the most recurring ligands, we found several that are either consistently, e.g. *ADM*, *CXCL8*, *CXCL5* and *GAL*, or prevalently, e.g. *CCL26,* associated with lower survival (Figure 4b; Table S6).

We reasoned that in those patients with higher expression of both receptors and ligands (or enzymes), a likely enhanced activation of the signaling axis should better correlate with specific patient phenotypes (e.g. lower or higher survival) than when considering individual components. We therefore assessed differences in the survival of patients grouped on the basis of the expression levels of GPCRs and ligand precursors (or enzymes) either alone or in combination. We then shortlisted receptor ligand pairs characterized by higher significance as well as greater hazard ratios of patients survival compared to the individual interaction partners (Figure 1a,4c; Table S8-9, see Methods). Intriguingly, we found several axes, such as *NPBWR2-NPW, PRHLR-NPY, PRLHR-PLR, CRHR2-UCN3, CALCR-CALCB,* associated with lower survival more significantly than considering interaction partners alone in multiple cancer types and (Figure 4c,d; Figure S4). Certain cancer types show multiple significant instances of associations between receptor-ligand axis and survival (e.g. lower survival in adrenal gland, eye; Figure 4c), suggesting a potential polypharmacology strategy for these tumor types.

Interestingly, it appears that some receptors participating to significant axes within the same cancer type share, to some extent, similar transduction mechanisms, suggesting that the patient survival phenotype might be contributed by multiple ligand stimuli converging on common signaling cascades. For example, in head and neck cancers we found significant correlation of higher survival and higher expression of *SSTR2-SST* (Figure S5a*); RXFP3-INSL5; RXFP1-RLN3* signaling axis, which also share similar coupling profiles (e.g. Gq/11, particularly *GNA14*, as well as Gi/o; Figure S6). *RXFP1-RLN3* is also associated with lower survival in breast carcinoma patients, highlighting the specificity of survival associations of these axes with cancer subtypes. Likewise, *SSTR2-SST* is associated to lower survival in Adrenal Gland patients (Figure S5b), *OPRK1-POMC* with lower survival in uveal melanoma and higher survival in Lining of body cavities, *CXCR2-CXCL6* with lower survival in Adrenal Gland and higher survival in Lining of body cavities, *TACR2-TAC1* is associated with lower survival in Liver cancer and higher survival in Cervix a1nd Uterus.

Notably, some of the receptor-ligand pairs that we found differentially co-regulated in terms of expression, are also significantly associated with survival differences (Figure 4e). In details,

we found a total of 21 receptor-ligand pairs from 14 cancer subtypes. *CCR5* is the receptor forming axes (with *CCL7* and *CCL8*) that are most frequently co-regulated and at the same time significantly associated with survival. In particular, *CCR5* co-up regulation is invariably associated with a better prognosis, irrespective of the cancer subtype considered. Other axes, such as those mediated by *CXCR2,* display cancer subtype-specific patterns of both differential expression regulation and correlation with patient survival (Figure 4e). Other axes, such as *CCR1-CCL4* are always co-up regulated but with different prognostic values, i.e. they are associated to lower survival in gastric cancer ("GI.GS" subtype) and to higher survival in melanoma (RAS and BRAF hotspot mutants subtypes, Figure 4e). For certain tumor types, such as skin melanoma or breast carcinoma, we found several subtype instances each characterized by different patterns of signaling axes being at the same time significantly co-regulated and associated to survival (Figure 4e).

Biosynthetic enzymes and cognate receptors form surrogate axes with prognostic value.

We found multiple ligand-synthesizing enzymes whose expression correlates with patients' survival in multiple cancers, with 55% of them significantly linked to worse survival prognosis (LogRank P-value < 0.01). Remarkably, the 5'-nucleotidase *NT5E* is invariably associated to lower survival in six distinct cancer types, including breast, stomach, pancreas, lung, eye and lining of body cavities (Figure 4b; Table S7). Several other enzymes appear to be always associated to lower survival, including: *PTGIS, ADCY3, NCAT2, ASS1, ADA, CHAT*.

Wherever the cognate ligand was a small organic molecule, we considered the corresponding synthesizing enzymes and we similarly checked for an improvement of the survival statistics when considering the combined expressions of interaction partners (Methods; Table S9). We found several instances involving *CHAT* and muscarinic receptors *CHRM1,3,5* associated with lower survival in multiple cancers (Fig. 5a,b). The three axes involving muscarinic receptors and *CHAT* are all found significant in liver cancer, with *CHRM3-CHAT* being recurrently correlated in three cancer types (Fig. 5a). We also found multiple instances of glutamate receptors (e.g. *GRM1,2,4,5,7*) and Glutamate decarboxylase 1 and 2 (*GAD1,2*). While the axes *GRM2-GAD1* shows tissue specific correlation with survival, being associated with lower survival in Lining of body cavities and with higher survival in Brain, the axes of *GAD2* with *GRM1,4,5* are always significantly associated with lower survival in Endometrium and Thyroid cancer tissues. We also found several instances of receptors for purines to multiple enzymes of the Purine catabolism pathway. In details, while the *ADORA2B-PNP* axes is invariably associated to lower survival in three distinct tissues (Lining of bodies cavity, Uterus, Brain) (Figure5a,b), *P2RY4* axes with *ENPP3* and *ENTPD1* is associated to higher survival in Uterus. The latter axes is in turn significantly associated with lower survival in

Adrenal Gland (Figure 5a). Likewise, *GPBAR1-SLC27A5* is associated to higher survival in Pancreatic cancer and Uterine cancer and to lower survival in Brain and White blood cell cancers.

Drugs for GPCR axes hold the potential to inhibits cancer cell growth

We retrieved GPCRs ligands showing cancer growth inhibitory capacity from a public resource containing the growth-inhibitory activity of 4,518 drugs tested across 578 human cancer cell lines[25]. A total of 13 out of 52 GPCR ligands were found to significantly inhibit the growth of cancer cell lines (Figure 6a)[25]. We checked the targets of these ligands on the pooled list of GPCRs that we found to mediate axes significantly associated with survival. Strikingly, all these drugs bind to receptors involved in significant axes with either ligands or enzymes (Figure 6a; Table S10). In particular, we found multiple ligands targeting certain GPCRs classes, including adenosine receptors (i.e. CGS-15943, MRS-1220, SCH-58261), muscarinic receptors (i.e. VU0238429, xanomeline, terfenadine) and opioid receptors (i.e. BNTX and JTC-801). Interestingly, the vast majority (76%) of these drugs are antagonists. Moreover, receptors targeted by antagonists appear to be more promiscuous, coupling primarily to $G_{q/11}$ or $G_{i/o}$, but also to a less extent to G_s and $G_{12/13}$ proteins (Figure 6c).

Discussion

In recent years the multifaceted role of GPCR signaling in cancer has begun to emerge[14,15], though an overall mechanistic picture is still lagging. At a molecular level, this is mainly due to the lack of a precise mechanistic understanding of the precise repertoire of transducers used by each receptor to transduce an external signals intracellularly, as well as how the upstream ligand and biosynthetic pathway components contribute to trigger an aberrant signaling. At a higher level, the complex circuitry wired by multi-component GPCRs systems is highly tailored to specific cellular contexts. Pathway dysregulation might happen through perturbations of the multi-component system at different levels, potentially converging to an overall dysregulation of the downstream signal.

In the present study, we compiled a GPCR centric interaction network, entailing both intracellular transduction pathways, as well as upstream ligands and biosynthetic pathways. We systematically inferred GPCR pathways activities from cancer transcriptomic datasets, with a special focus on the dysregulation of GPCR signaling axes involving their extracellular components.

We found that multiple GPCRs are co-regulated together with their cognate ligands or enzymes, suggesting a coordinated effect on the signaling axis. Some of these are widespread co-regulated across cancers, others are affected in specific cancer contexts. Chemokine receptors and cognate chemokines are certainly the most co-regulated axes. The precise role in pro- or anti-tumor immunity played by these chemokine receptor axes is known to depend on the precise cellular and tumor subtype context[14,26]. For example, the known tumor-promoting *CXCR2* is found co-down regulated together with *CXCL2* in liver cancer (iCluster2 subtype) and with *PPBP*, the precursor of *CXCL7*, in lung (LUAD.3 subtype). In both cases, the co-down-regulation is associated with higher survival (Figure 4e). Likewise, co-up regulation of the latter axis is associated with lower survival in stomach cancer (GI.GS subtype). Other chemokine receptors, such as *CCR1, CCR3* or *CCR5*, are significantly co-up regulated across different cancer subtypes. While *CCR1-CCL4* has distinct correlation to survival depending on the cancer type considered, *CCR5-CCL7* or *CCR5-CCL8* co-up regulation is always associated with higher survival, suggesting a general anti-tumor immunity of these axes in the cancer subtypes considered (Figure 4e).

Interestingly, we also found a few non-chemokine receptors that are widespread co-up activated, such as *KISS1, NPBWR, MCHR1 or APLNR* which suggest signaling axis mechanisms other than chemokine ones sustaining the interaction with the TME. Another notable example of highly context specific co-regulation of signaling axes is the one mediated by *CALCR* and its close paralogue, *CALCRL*, whose interaction with *ADM* is a critical factor determining relapse and drug resistance in Leukemia[27].

We found that many of the co-regulated signaling axes are also associated with patient survival. For example, in head and neck cancer we have found multiple signaling axes significantly associated to higher survival, such as *SSTR2-SST*, *RXFP3-INSL5* and *RXFP1-RLN3*. Intriguingly, all the three receptors show converging couplings towards $G_{q/11}$ and Gi/o proteins. Notably the *SSTR2-SST* has been significantly associated to higher survival rates in EBV-nasopharyngeal cancer, a type of head and neck cancer, and its targeting with PEN-221 has been proposed to treat this type of tumor[28]. *SSTR2-SST* is also one of the axes with the strongest literature support to cancer, along with several others including e.g. *MC1R-POMC*, *GHRHR-GHRH* (Figure S7; Table S11).

Along with the mRNA expression of peptide ligand precursors, we also considered the expression of synthesizing enzymes as a proxy to model upstream signals from small organic ligands. Indeed we found a significant correlation of patients' survival with expression of axes formed by receptors and biosynthetic enzymes. Similarly to receptor-ligand axes, we found

enzymatic axes recurrently associated with survival in a cancer-type specific fashion. For example, we found multiple instances of muscarinic receptors and Choline O-acetyltransferase (*CHAT*). Cholinergic signaling has been established as a key factor mediating gastric cancer tumorigenesis via an NGF promoted feed-forward loop[18]. Intriguingly, the association between the expression of muscarinic receptors-*CHAT* axis with lower survival in hepatocellular carcinoma correlates with the known involvement of acetycholinesterase, which opposes *CHAT* function by degrading acetylcholine, and is associated to better prognosis in HCC[29]. Other biosynthetic enzyme axes that we have identified are also supported by additional lines of evidence from the literature (Figure S7d). For example, the *CNR1-FAAH* axis, which is associated to higher survival in pancreatic cancer[30]. While in this study we considered either whole biosynthetic pathways or individual enzymes producing a given ligand, future efforts will be directed to also consider rate-limiting enzymes as a proxy for ligands abundances.

This information is indeed particularly important for receptor systems such as GPCRs, which frequently bind to organic ligands, and might represent a resource complementing existing knowledge-bases for receptor-ligand interactions used to characterize cell-cell communication from spatial transcriptomics datasets[20]. It will be critical in the future to dissect the precise cell context-specific role played by these GPCRs signaling axes by analyzing single cell as well as spatial transcriptomics RNAseq datasets.

Strikingly, all of the GPCR drugs that have been reported to significantly hamper cancer cell line growth from a recent systematic screen[25] target receptors forming axes with either ligands or enzymes found significantly correlated with patient survival in our study. This suggests that targeting these GPCRs axes could directly hamper cancer growth. Enzymes can also constitute additional therapeutic targets, either through small molecules inhibition or antibodies. Indeed, antibodies anti-*NT5E* and anti-*ENTPD1* are currently being evaluated in combination with cancer immunotherapies.

The potential of repurposing GPCR drugs to be used in combination with (immuno-) therapies shall be more extensively explored in the future.

Methods

GPCR signaling network

To create an extended GPCR network, we have extracted the GPCR and ligand interactions from the IUPHAR/BPS Guide to PHARMACOLOGY database (IUPHAR database)[31]. We

retained only endogenous ligands, thereby extracting 243 GPCRs and 328 ligands. Based on database classification, the considered ligands can be grouped into five groups: 211 peptides, 107 metabolites, 7 synthetic organic, 2 inorganic and 1 natural product. All the other annotations describing GPCRs, and ligands were also pulled from the database. To further extend our network we added enzymes that produce metabolite ligands. To this end we used Rhea, a curated database of chemical reactions of biological relevance[32]. To identify metabolizing enzymes of GPCRs ligands based, we first referenced IUPHAR ligands to ChEBI, a collection of known 'small' chemical entities of biological interest[33], which is used by Rhea to identify ligands. Ligand mapping between IUPHAR and Rhea was carried out by using the IUPHAR International Chemical Identifiers (InChis). As many of the chemical reactions are happening under physiological pH of 7.3, the ChEBI identities of participants in Rhea reactions at pH 7.3 were considered. Using our in-house python scripts, we have extracted all the enzymes catalyzing reactions where GPCR ligands participate as products, meaning that most of the reactions have left-right (LR) directionality or, in some cases, opposite (right-left - RL) direction. Moreover, we also considered all the enzymes catalysing reactions of unknown (UN) direction. Since it is possible that enzyme annotation is available only for parent reactions, i.e., more generic chemical reactions, we also added parent reactions for each of the already extracted reactions based on Rhea's hierarchical reaction relationships. For each of the Rhea identity acquired, we extracted the corresponding human enzyme whenever available. We have extracted 1492 enzymes linked to our ligands. We have further filtered out enzymes classified as kinases, channels, membrane-bound or involved in transport based on Pfam[34], Panther classification[35], Gene ontology processes (GO)[36] or full UniProt[37] enzyme name. After this filtering step, 872 enzymes remained which were further filtered out by using the STRING database, a knowledgebase of known or predicted protein-protein functional interactions[38]. Via STRING, we mapped all possible interactions between GPCRs and enzymes producing cognate ligands, by using a minimum confidence score of 150, and leading to a shortlist of 128 enzymes. The final network contained 128 enzymes, 328 ligands and 243 receptors, for a total of 182 binary enzymes-ligand interactions and 646 ligand-GPCR interactions. The whole network was compiled containing different descriptions of each of the three main interactors and deposited to SIGNOR database[39].

We generated G protein specific gene set as well as annotations by gathering consensus experimental transducer data (i.e. Universal Coupling Map[9]) as well as predicted through Precogx[7].

Transcriptomics Dataset

We considered gene-level expressions from the dataset 'TCGA TARGET GTEx' publicly available at UCSC Xena (https://xenabrowser.net/datapages/)[40]. These data were generated using the TOIL pipeline, a portable, open-source workflow software that can be used to run scientific workflows on a large scale in cloud or high-performance computing (HPC) environments[41]. The Toil pipeline uses STAR[42] to generate alignments and read coverage graphs, and performs quantification using RSEM[43] and Kallisto[44]. It was utilized to process ~20,000 RNA-seq samples to create a consistent meta-analysis of datasets free of computational batch effects.

We extracted the gene-level expression data, *RSEM expected_count*, using the R package 'UCSCXenatools' for 19109 samples. Thereafter, the TARGET samples were filtered and only TCGA and GTEx samples were retained (n=18305; Table S2). Due to unavailability of GTEx data for all the tissue types, only 16 whole cancer-types (TCGA) were contrasted to the healthy samples pertaining to the same tissue (GTEx). The information for 76 distinct TCGA subtypes was retrieved from TCGAbiolinks[45] (R4.2, Bioconductor v3.16) except for 'Pancreas' and 'Testis' for whom it was unavailable. The contrast for each subtype was performed using the corresponding GTEx whole tissue dataset.

Differential Expression Analysis

We obtained log-transformed gene-level count data (RSEM expected_count) from the UCSCXenatools[40], which was subsequently inverse-transformed to raw expected count. TCGA matched normal samples were discarded, and a differential expression (DE) analysis was performed to identify genes with significant expression changes across TCGA-GTEx samples. The DE analysis was conducted using DESeq2[46], a widely used pipeline for DE analysis. We considered genes with an |LFC|>1 and an adjusted p-value (Padj) < 0.01 (Benjamini-Hochberg correction[47]) to be differentially expressed. This process was performed for each of the 16 whole tissue types and 76 subtypes.

Co-regulation Analysis

To identify co-differentially regulated receptor-ligand pairs, we assigned scores to each of these gene pairs. We considered three major cases: (a) both co-recurrently up/downregulated with significant Padj values (a score of -2/+2), (b) both co-recurrently up/downregulated with at least one of them having a significant Padj value (a score of -1/+1), and (c) no or anti co-differential regulation (a score of 0). This information was displayed in the form of Seaborn clustermap (v0.11.2, Python v3.9.7) with hierarchical clustering performed using the 'ward' method of the scipy.cluster library (v1.7.1, Python v3.9.7). On top of this heatmap, G-coupling information (ref Methods section GPCR - G Protein coupling data) and ligand's mechanism of

action (from IUPHAR) was annotated as separate heat maps using matploltib library (v3.4.3, , Python v3.9.7) This way we identified differentially expressed genes and co-differentially regulated receptor-ligand pairs in a comprehensive and robust manner. The use of DESeq2 and Benjamini-Hochberg correction ensured the statistical significance of the results, and the consideration of multiple tissue types and subtypes allowed for a more nuanced understanding of the data.

Gene-Set Enrichment Analysis (GSEA)

We performed gene set enrichment analysis (GSEA) using the WEB-based GEne SeT AnaLysis Toolkit (WebGestalt)[48]. The GSEA process was automated using the R package WebGestaltR. We chose the 'reactome' pathway database[49] for enrichment analysis. The results from these analyses were employed to correlate the DE genes to various reactome pathway gene sets with a positive (enriched in TCGA) or negative enrichment (enriched in GTEx) score with Padj<0.01. We considered GSEA results for downstream analyses by filtering pathways containing the GPCR ligands or enzymes belonging to either Signal transduction ('signaling by GPCR':R-HSA-372790) or Metabolism ('metabolism':R-HSA-1430728) reactome upper level domains.

Comparative Transcriptome-Metabolome analysis

We mapped differentially expressed transcripts and metabolites onto a functional interaction network from Reactome pathways[5051], which integrates in the same network information from proteins and small organic molecules. We first collected data from the Pan-Cancer Metabolism Data Explorer[21], a comprehensive database that contains information about the differential regulation of metabolites associated with cancer aggressiveness. This resource gathered data from over 900 patient samples from five subtypes of cancer: Breast (BRCA), Prostate (PRAD), Pancreas (PAAD), Kidney (KIRC) and Bladder (BLCA). Using this data, we identified the differential expression regulation of enzyme(s) associated with ligand biosynthesis. For this purpose, we used the list of our curated enzymes to query within this external dataset using an HMDB to ChEBI mapping done via MetaboAnalyst[52]. Then, for each pathway, the LFC and Padj values of constituent enzymes were retrieved. Subsequently, this information was merged with the reactome functional interaction network for a ligand-receptor signaling pathway to complete the entire axis. This was achieved by considering only the first neighbors for involved participants. The resultant networks were used to visualize the directional regulation as well as the DE profiles by utilizing Cytoscape[53].

Pathway Abundance Score

We compared the enrichment scores of differentially expressed (DE) transcripts with the differential abundance scores of DE metabolites obtained from the Pan-Cancer Metabolism Data Explorer. To this end, we respectively calculated the GSEA Enrichment Score (ES) and the Pathway Abundance Score (PAS)[21] by considering only the filtered set of curated biosynthetic pathways from Reactome, specifically for metabolism (R-HSA-1430728). The PAS was calculated by using the formula (I-D)/S, where I represents the number of metabolites (or enzymes) upregulated (LFC>0) in a pathway, D represents the number of metabolites downregulated (LFC<0) in the pathway, and S is the sum of I and D or 1 if I+D=0. A pathway is considered enriched in cancer if the PAS is positive and negative otherwise (similar to ES). We utilized these two metrics to qualitatively evaluate the concordance of enriched biosynthetic pathways in our findings.

Survival Analysis

The study used univariate unadjusted Cox-Proportional Hazard (Cox-PH) regression models to compute hazard ratios (HR) for predicting the risks of death between high-risk and low-risk TCGA patient groups based on overall survival time. The survival and censoring information was retrieved using TCGAbiolinks for all TCGA tissues. The risk groups were stratified based on median values of expression, and Kaplan-Meier (KM) plots were used to compare their survival curves. 'survival' (v4.3.0) and 'survminer' packages (v0.4.9) in R (v4.2) were used for survival analyses, and statistical significance between survival curves was estimated using log-rank tests with p-values less than 0.05 considered significant.

For the individual components (i.e. Receptors/Ligands/Enzymes), patients were stratified in two risk groups based on the median expression value of each component and HR were estimated. Using the ‘low expression’ group as reference for HR calculations, we label the resulting risk groups as ‘High Survival’ (HR>1) or ‘Low Survival’ (0<HR<1). We further employed the combined expression levels for GPCR axes components i.e. GPCR-ligand/GPCR-enzyme for risk stratification. The patients were now stratified in the two risk groups utilizing the median cutoffs of both the components for e.g. for a GPCR-Ligand based stratification, patients with expression value, EV>MedianGPCR and EV>MedianLigand were placed in one group and the patients with EV<MedianGPCR and EV<MedianLigand in another group. We then filtered out those significant GPCR-axis pairs wherein the Hazard-Ratio axis (HR_{axis}) was more than two-fold of the HR in both the individual stratifications (i.e. for above

example $0.5<HR_{axis}/HR_{GPCR}<2$ and $0.5<HR_{axis}/HR_{Ligand}<2$) along-with the criteria that the logrank-p_{axis} is smaller than the logrank-p values value of both of the individual stratifications. These instances were then displayed on a heatmap created using Seaborn clustermap (v0.11.2, Python v3.9.7) with hierarchical clustering performed using the 'ward' method of the scipy.cluster library (v1.7.1, Python v3.9.7).
Survival analysis were performed considering whole cancer tissues, as defined in Xena browser, or at the TCGA molecular subtype level.

Analysis of GPCR drugs on PRISM drug repurposing resource

We retrieved GPCR ligands with cancer growth-inhibitory capacity from a publicly available resource PRISM (Program for Response to Inter- and Intra-Species Variation in Cancer Pharmacology)[25]. This drug repurposing resource was developed to identify potential drugs for cancer treatment. To predict drug response, PRISM utilizes genomic data from cancer cell lines and patient-derived xenografts, which are inputted into machine learning pipelines. By taking into account inter- and intra-species variability in drug response, PRISM aims to provide personalized treatment options for specific cancer subtypes. The dataset retrieved from PRISM contained information on the growth-inhibitory activity of over 4,500 drugs tested on 578 human cancer cell lines. This data was filtered based on the drugs which had GPCRs as targets and a list of 52 GPCR drugs/ligands was obtained. To identify active drugs for cancer treatment, PRISM uses the bimodality coefficient and correlation as metrics. The bimodality coefficient measures the separation between the two peaks in a drug response distribution, identifying drugs with a bimodal distribution of response across cancer cell lines. The correlation measures the strength of the relationship between predicted and actual drug response in patient-derived xenograft models. Following a similar approach, we used these metrics (correlation>0.2, bimodality coefficient>0.35) to identify GPCR drugs that are active against specific cancer subtypes. The active drugs were then investigated on the basis of GPCRs associated with survival in either of GPCR-Ligand or GPCR-Enzyme axes. The information was further annotated with G-coupling information and ultimately displayed using visual representations such as scatter plots, venn diagram and heatmaps utilizing seaborn (v0.11.2, Python v3.9.7) and matplotlib libraries (v3.4.3, , Python v3.9.7).

Availability

Code and data used for this study are available at: https://gpcrcanceraxes.bioinfolab.sns.it/

Conflict of interest

None declared

Funding

F.R. was supported by the Italian Ministry of University and Research through the Department of excellence “Faculty of Sciences” of Scuola Normale Superiore. The research leading to these results also received funding from the Italian Association for Cancer Research (AIRC) under My First AIRC Grant (MFAG) 2020 - ID. 24317 project – P.I. F.R; the US National Institutes of Health (U24 HG012198) to G.W.

Acknowledgments

We gratefully acknowledge the CINECA award, in collaboration with AIRC, for the availability of high performance computing resources and support. We gratefully acknowledge computational resources of the Center for High Performance Computing (CHPC) at Scuola Normale Superiore.

Figure Legends

Figure 1 Workflow and overview of the GPCR system and dataset considered a) The workflow of the study illustrating the datasets employed as well as the main downstream analysis; b) Sunburst charts providing an overview of the number of liganded and orphan G-protein-coupled receptors (GPCRs) considered; Classification of the liganded GPCRs based on their ligand type, i.e. either peptide, organic, or a combination of both; number of enzymes that are either currently included or absent within the Reactome pathways. The included ligands are additionally subcategorized based on the frequency of their types and the frequency of the Reactome pathway domains with which they are linked. For enzymes, only the second layer of this information is shown; c) Bar plot exhibiting the frequency of differentially expressed (TCGA vs. GTEx) GPCRs, ligands, and enzymes corresponding to each whole tissue. d) Stacked bar-plot displaying the number of significantly enriched ligand-associated pathways within Reactome sub-domains: Signaling ('R-HSA-372790: signaling by GPCR') and Metabolism ('R-HSA-1430728:metabolism'); e) heatmap showing ligand-associated pathways which are significantly enriched in at least one tissue.

Figure 2 Peptide Ligand-Receptor co-regulation a) Heatmap (center panel) displaying the co-differential regulation for Receptor-Ligand pairs across different TCGA subtypes (color-coded at the top row). Darker red represents both receptor and ligand significantly co-up regulated in TCGA (i.e. LFC>1 and Padj<0.01) and darker blue represents both receptor and ligand significantly co-down regulated in TCGA (i.e. LFC<1 and Padj<0.01). Paler colors represent either of the receptor-ligand as significantly DE (i.e. |LFC|>1 for both but Padj<0.01 for only one of these). White cells indicate anti-regulation or no fold change at all in at least one of them. Only pairs affected in at least 25% of TCGA subtypes are displayed. A dashed line separating the two clusters, created using Hierarchical clustering, is shown in the middle. Heatmap (left panel) uses color codes to display the ligand's mechanism of action in 'Action', G Protein-coupling associated with the GPCRs and GPCR Family. Heatmap (right panel) displays the information contained in the left panel as concise bar plots for each of the two clusters. For each cluster barplots indicate G-protein coupling preferences as well as GPCR families; b) The scatter plot displays a GPCR-centric view of the heatmap in (a) based on tissue-wise or ligand-wise similarity. Here, ligand-wise similarity for each GPCR is the average of the pairwise Euclidean distances between co-differentially expressed (co-DE) profiles across its ligands (i.e. vertically), while tissue-wise similarity is a similar metric across tissues (i.e. horizontally). The size of the markers is proportional to the number of ligands, and the colors correspond to the average difference in the number of co-up and co-down profiles. GPCRs located in the top-right corner are more diverse in terms of their ligand interactions; c) Diverse co-DE profiles are observed with respect to different ligands and tissues corresponding to the same GPCR. For example, CXCR3 and its corresponding ligands exhibit an almost exclusively co-upregulated profile irrespective of tissue type, while CALCRL and its ligands display varied co-regulation profiles, implying a varying interaction with its ligands in different tissue (or cancer) types.

Figure 3: **Biosynthetic pathway enrichment** (a) Heatmap of GSEA pathway enrichment analysis for differentially expressed (DE) genes in cancer (TCGA) over normal (GTEx) samples, considering Reactome pathways related to signaling and metabolism. Red cells indicate pathways enriched in TCGA, blue cells indicate pathways enriched in GTEx, and

white cells indicate no enrichment. Cells marked with '*' correspond to a statistically significant enrichment (Padj < 0.01).
(b) Concordant pathway enrichment analysis considering both DE genes (Enrichment Score, ES) and metabolites (Pathway Abundance Score, PAS). A score of 1 (red) indicates co enrichment in cancer (ES & PAS > 0), and a score of -1 (blue) indicates co enrichment in normal samples (ES & PAS < 0).
(c) Bar plot displaying numerical values for ES and PAS for the 'Purine catabolism' pathway in three different tissues: Breast, Bladder, and Prostate.
(d) Heatmap of DE genes for the 'Purine catabolism' pathway in Breast cancer. Red cells indicate upregulated genes in TCGA, blue cells indicate downregulated genes in TCGA, and cells marked with '*' correspond to a statistically significant differential expression (Padj < 0.01).
(e) A functional interaction network between genes (ovals) and metabolites (diamonds) in the 'Purine catabolism' pathway in Breast cancer. Red nodes indicate upregulated components, blue nodes indicate downregulated components, and green nodes indicate no information available. The network shows that over-activation of the pathway is contributed by over-expression of both genes and metabolites.

Figure 4 Association of GPCR-peptide axes to survival (a) Survival association of G protein-coupled receptor (GPCR) components in various cancer types. The funnel plot shows the number of significant instances of individual components, such as receptors, ligands, and enzymes, whose expression values are associated with patient survival across various cancer types. A total of 298 unique significant receptor instances (log rank p-value < 0.05) were identified in 22 cancer tissues, 108 ligands in 23 cancers, and 99 enzymes in 17 cancers. (b) Frequency bar-plots for top 20 individual instances of receptors, enzymes and ligands associated with patient survival in various cancer types. c) The bubble plot shows the correlation between the combined-expression levels of GPCR-ligand pairs and patient survival across various cancer types. The pairs with a log rank p-value <0.05 and also lower than log rank p-values for individual GPCR/ligand are displayed. Bubble color is proportional to HR: i.e. HR>1 : High expression is correlated with high survival (red); HR<1: High expression is correlated with poor survival (blue). Bubble diameters is proportional to the -log(log-rank p-value). (d) The expression values for the CALCR-CALCB axes in breast cancer (right) and corresponding Kaplan-Meier curve for survival analysis of patients stratified based on the individual as well as combined receptor/ligand expression;(e) Scatter-plot showing with differentially co-expressed GPCR-ligand pairs (upper triangle=co-up regulation; lower triangle=co-down regulation) that were significantly associated with patient survival (red=higher survival; blue=lower survival). A total of 21 receptor-ligand pairs from 14 cancer subtypes were identified.

Figure 5 (a) The bubble plot shows the correlation between the combined-expression levels of GPCR-enzyme pairs and patient survival across various cancer types. The pairs with a log rank p-value <0.05 and also lower than log rank p-values for individual GPCR/enzyme are displayed. Bubble color is proportional to HR: i.e. HR>1 : High expression is correlated with high survival (red); HR<1: High expression is correlated with poor survival (blue). Bubble diameters is proportional to the -log(log-rank p-value). (b) The CHRM3-CHAT axis is recurrently correlated with lower survival rates in three types of cancer: Liver, Brain and Esophagus. The KM plots (left) display the risk enhancement achieved when CHRM3-CHAT combined stratification was applied as compared to individual CHRM3/CHAT as evident by

log rank p-values. Similarly, the ADORA2B-PNP axis is also consistently associated with lower survival rates in three different types of tissues: Brain, Uterus and Lining of body cavities. The KM plots on the right show the improved stratification due to ADORA2B-PNP pair in Brain cancer.

Figure 6 GPCR ligands with cancer growth-inhibitory capacity (a) The scatterplot displays 52 GPCR ligands that were tested in 578 cell lines in PRISM drug repurposing resource.13 GPCR ligands (in red) were found to significantly inhibit the growth of cancer cell lines (correlation>0.2, bimodality coefficient>0.35). (b) Venn diagram showing the overlap of GPCRs in the target proteins of these 13 ligands with the pooled list of GPCRs that were found to mediate survival associated GPCR-ligand/GPCR-enzyme axes. (c) The heatmap displays the 13 shortlisted ligands and their corresponding GPCR targets. The targets marked with '*' are the overlapping GPCRs with significant axes. Heatmap also displays the overall action mode of ligand, more detailed mechanism of action (MOA) and G-protein coupling associated with the GPCR.

References

1. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol* **3**, 639–50 (2002).
2. Oldham, W. M. & Hamm, H. E. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol* **9**, 60–71 (2008).
3. Wootten, D., Christopoulos, A., Marti-Solano, M., Babu, M. M. & Sexton, P. M. Mechanisms of signalling and biased agonism in G protein-coupled receptors. *Nat Rev Mol Cell Biol* 1 (2018) doi:10.1038/s41580-018-0049-3.
4. de Mendoza, A., Sebé-Pedrós, A. & Ruiz-Trillo, I. The evolution of the GPCR signaling system in eukaryotes: modularity, conservation, and the transition to metazoan multicellularity. *Genome Biol Evol* **6**, 606–619 (2014).
5. Flock, T. *et al.* Selectivity determinants of GPCR-G-protein binding. *Nature* **545**, 317–322 (2017).
6. Inoue, A. *et al.* Illuminating G-Protein-Coupling Selectivity of GPCRs. *Cell* **177**, 1933-1947.e25 (2019).
7. Matic, M. *et al.* PRECOGx: exploring GPCR signaling mechanisms with deep protein representations. *Nucleic Acids Res* **50**, W598–W610 (2022).
8. Singh, G., Inoue, A., Gutkind, J. S., Russell, R. B. & Raimondi, F. PRECOG: PREdicting COupling probabilities of G-protein coupled receptors. *Nucleic Acids Res* **47**, W395–W401 (2019).
9. Hauser, A. S. *et al.* Common coupling map advances GPCR-G protein selectivity. *Elife* **11**, (2022).

10. Avet, C. *et al.* Effector membrane translocation biosensors reveal G protein and βarrestin coupling profiles of 100 therapeutically relevant GPCRs. *Elife* **11**, 1–34 (2022).
11. Foster, S. R. *et al.* Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors. *Cell* **179**, 895-908.e21 (2019).
12. O'Hayre, M. *et al.* The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. *Nat Rev Cancer* (2013) doi:10.1038/nrc3521.
13. Ramms, D. J. *et al.* Gαs–Protein Kinase A (PKA) Pathway Signalopathies: The Emerging Genetic Landscape and Therapeutic Potential of Human Diseases Driven by Aberrant Gαs-PKA Signaling. *Pharmacol Rev* **73**, 155–197 (2021).
14. Wu, V. *et al.* Illuminating the Onco-GPCRome: Novel G protein-coupled receptor-driven oncocrine networks and targets for cancer immunotherapy. *J Biol Chem* (2019) doi:10.1074/jbc.REV119.005601.
15. Arang, N. & Gutkind, J. S. G Protein-Coupled receptors and heterotrimeric G proteins as cancer drivers. *FEBS Lett* **594**, 4201–4232 (2020).
16. Raimondi, F. *et al.* Rare, functional, somatic variants in gene families linked to cancer genes: GPCR signaling as a paradigm. *Oncogene* **38**, 6491–6506 (2019).
17. Renz, B. W. *et al.* β2 Adrenergic-Neurotrophin Feedforward Loop Promotes Pancreatic Cancer. *Cancer Cell* **33**, 75-90.e7 (2018).
18. Hayakawa, Y. *et al.* Tumorigenesis through Aberrant Cholinergic Article Nerve Growth Factor Promotes Gastric Tumorigenesis through Aberrant Cholinergic Signaling. *Cancer Cell* **31**, 21–34 (2017).
19. Armingol, E., Officer, A., Harismendy, O. & Lewis, N. E. Deciphering cell-cell interactions and communication from gene expression. *Nat Rev Genet* **22**, 71–88 (2021).
20. Dimitrov, D. *et al.* Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat Commun* **13**, (2022).
21. Reznik, E. *et al.* A Landscape of Metabolic Variation across Tumor Types. *Cell Syst* **6**, 301-313.e3 (2018).
22. Patra, K. C. *et al.* Mutant GNAS drives pancreatic tumourigenesis by inducing PKA-mediated SIK suppression and reprogramming lipid metabolism. *Nat Cell Biol* **20**, 811–822 (2018).
23. Ideno, N. *et al.* GNASR201C Induces Pancreatic Cystic Neoplasms in Mice That Express Activated KRAS by Inhibiting YAP1 Signaling. *Gastroenterology* (2018) doi:10.1053/J.GASTRO.2018.08.006.
24. Cekic, C. & Linden, J. Adenosine A2A receptors intrinsically regulate CD8+ T cells in the tumor microenvironment. *Cancer Res* **74**, 7239–7249 (2014).
25. Corsello, S. M. *et al.* Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat Cancer* **1**, 235–248 (2020).
26. Ozga, A. J., Chow, M. T. & Luster, A. D. Chemokines and the immune response to cancer. *Immunity* **54**, 859–874 (2021).
27. Larrue, C. *et al.* Adrenomedullin-CALCRL axis controls relapse-initiating drug tolerant acute myeloid leukemia cells. *Nat Commun* **12**, 1–15 (2021).
28. Lechner, M. *et al.* Somatostatin receptor 2 expression in nasopharyngeal cancer is induced by Epstein Barr virus infection: impact on prognosis, imaging and therapy. *Nat Commun* **12**, (2021).

29. Zhao, Y. *et al.* Acetylcholinesterase, a key prognostic predictor for hepatocellular carcinoma, suppresses cell growth and induces chemosensitization. *Hepatology* **53**, 493–503 (2011).
30. Michalski, C. W. *et al.* Cannabinoids in pancreatic cancer: correlation with survival and pain. *Int J Cancer* **122**, 742–750 (2008).
31. Harding, S. D. *et al.* The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res* **46**, D1091–D1106 (2018).
32. Lombardot, T. *et al.* Updates in Rhea: SPARQLing biochemical reaction data. *Nucleic Acids Res* **47**, D596–D600 (2019).
33. Hastings, J. *et al.* ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res* **44**, D1214–D1219 (2016).
34. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412–D419 (2021).
35. Mi, H. *et al.* PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* **49**, D394–D403 (2021).
36. Blake, J. A. *et al.* Gene ontology consortium: Going forward. *Nucleic Acids Res* **43**, D1049–D1056 (2015).
37. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **45**, D158–D169 (2017).
38. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607–D613 (2019).
39. Lo Surdo, P. *et al.* SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update. *Nucleic Acids Res* **51**, D631–D637 (2023).
40. Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* **38**, 675–678 (2020).
41. Vivian, J. *et al.* Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol* **35**, 314–316 (2017).
42. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
43. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, (2011).
44. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525–527 (2016).
45. Colaprico, A. *et al.* TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* **44**, e71 (2016).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
47. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* vol. 57 289–300 Preprint at https://doi.org/10.2307/2346101 (1995).
48. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res* **41**, (2013).
49. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res* **48**, D498–D503 (2020).

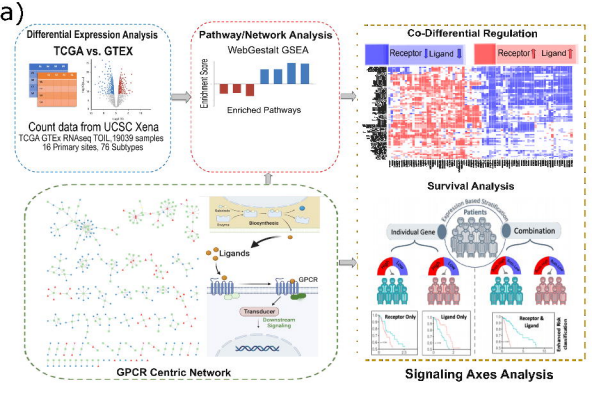
50. Wu, G., Feng, X. & Stein, L. A human functional protein interaction network and its application to cancer data analysis. *Genome Biol* **11**, R53 (2010).
51. Raimondi, F., Burkhart, J. G., Betts, M. J., Russell, R. B. & Wu, G. Leveraging biochemical reactions to unravel functional impacts of cancer somatic variants affecting protein interaction interfaces. *F1000Research 2021 10:1111* **10**, 1111 (2021).
52. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res* **37**, (2009).
53. Lotia, S., Montojo, J., Dong, Y., Bader, G. D. & Pico, A. R. Cytoscape app store. *Bioinformatics* **29**, 1350–1351 (2013).

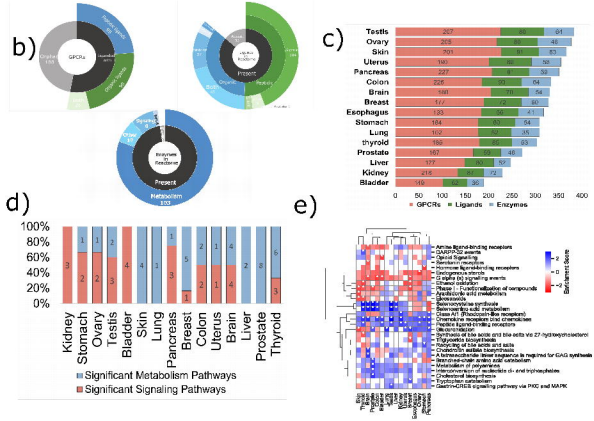
Supplementary Figures

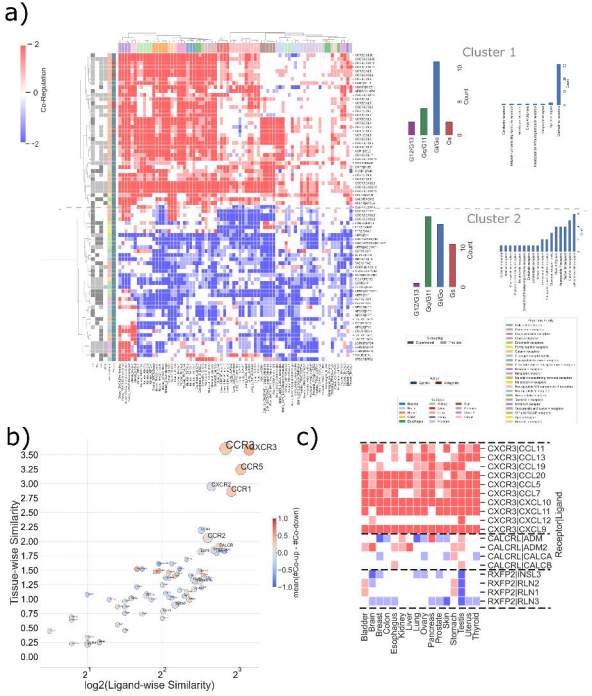
Figure S1: a) stacked barplot with the distribution of DE GPCRs for each cancer tissue categorized by coupling information; b) heatmap showing the LFCs from DE analysis (TCGA vs GTEX, via DESEQ2) of G protein in each cancer tissue. Significant instances are annotated with "*"; c) same as b) but for TCGA molecular subtypes;
Figure S2: Heatmap displaying the co-differential regulation for Receptor-Ligand pairs across different TCGA subtypes (color-coded at the top row). Darker red represents both receptor and ligand significantly co-up regulated in TCGA (i.e. LFC>1 and Padj<0.01) and darker blue represents both receptor and ligand significantly co-down regulated in TCGA (i.e. LFC<1 and Padj<0.01). Paler colors represent either of the receptor-ligand as significantly DE (i.e. |LFC|>1 for both but Padj<0.01 for only one of these). White cells indicate anti-regulation or no fold change at all in at least one of them.
Figure S3: functional interaction network between genes (ovals) and metabolites (diamonds) in the 'Purine catabolism' pathway in a) pancreatic and b) prostatic cancer. Red nodes indicate upregulated components, blue nodes indicate downregulated components, and green nodes indicate no information available. The network shows that over-activation of the pathway is contributed by over-expression of both genes and metabolites.
Figure S4:Kaplan-Meier curve for survival analysis done by stratifying patients based on expression values of representative GPCR-ligand axes, found to be more significantly associated with lower survival than individual components in several cancer types.
Figure S5: Kaplan-Meier curve for survival analysis done by stratifying patients based on expression values of SSTR2-SST axis in either head and neck or adrenal gland cancers.
Figure S6:G protein coupling preferences of axes found significantly associated with higher survival in head and neck: i.e. *SSTR2-SST*, *RXFP3-INSL5* and *RXFP1-RLN3*
Figure S7: number of Pubmed evidences obtained by querying with the following keywords: a) GPCR name *and* ligand name *and* cancer;b) GPCR name *and* ligand name *and* cancer *and* tissue; c) GPCR name *and* ligand name *and* cancer *and* immunotherapy; d) GPCR name *and* enzyme name *and* cancer

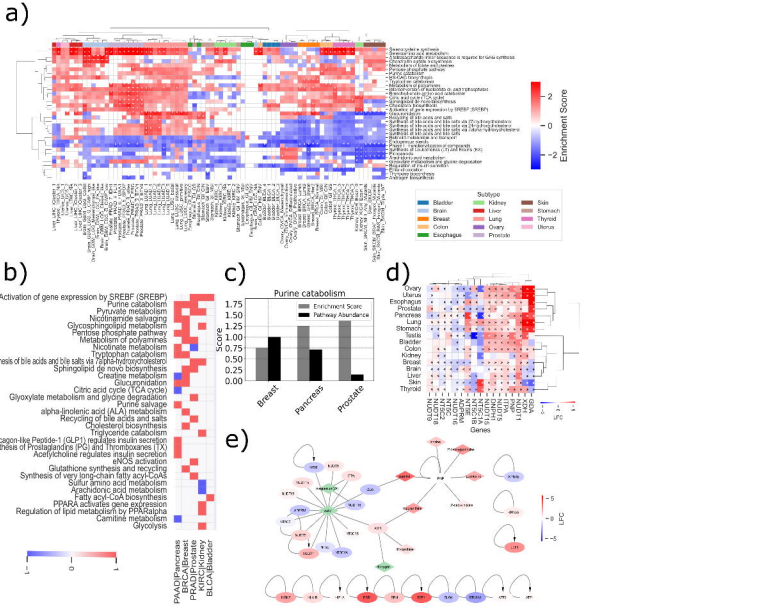
Supplementary Tables

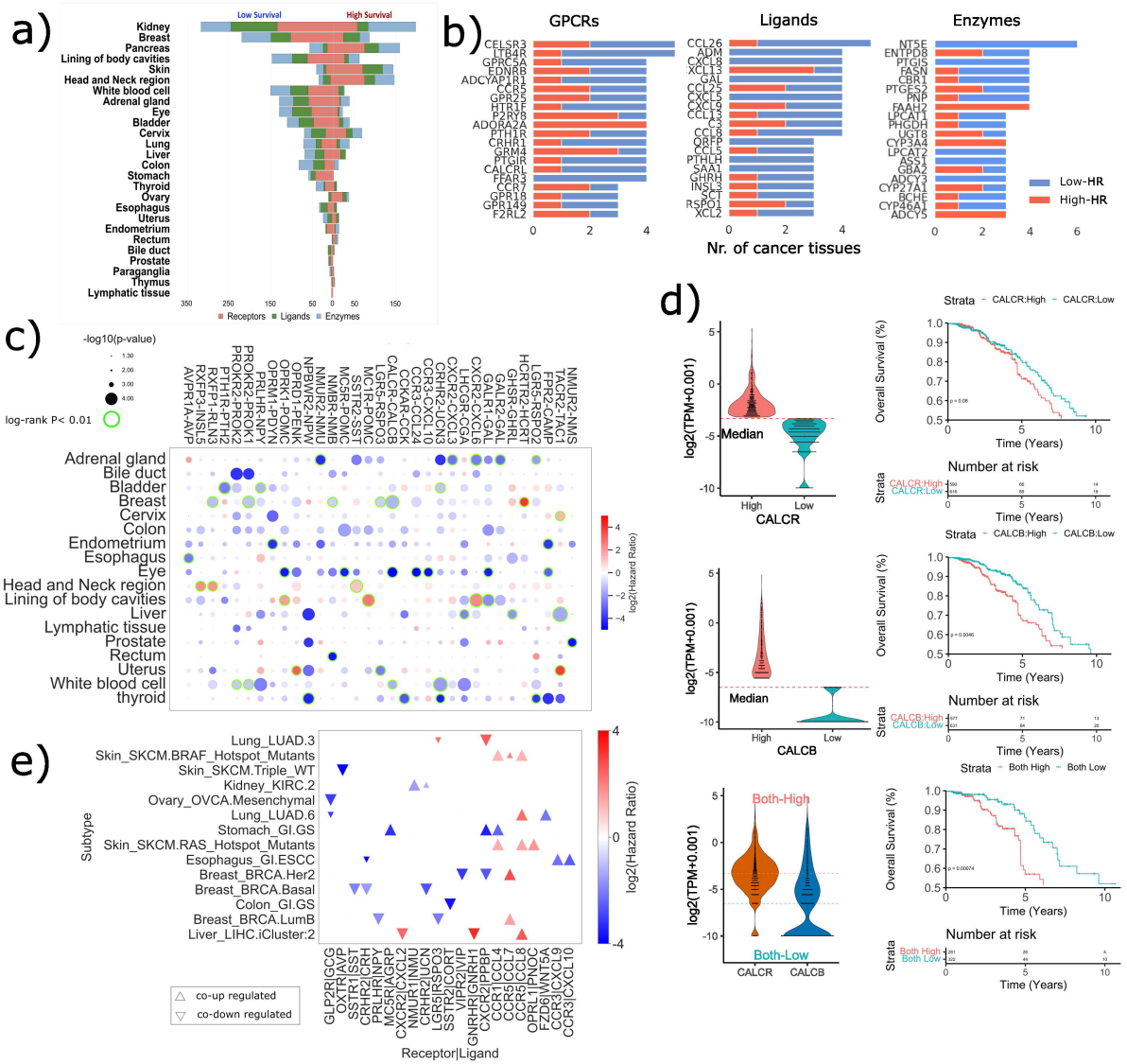
Table S1: GPCR ligands in Reactome Pathways (Metabolism and Signal Transduction domains)
Table S2. List of biosynthetic enzymes from RHEADB A1producing GPCR ligands
Table S3. Tissue-wise sample distribution in the data retrieved from UCSCXena.
Table S4. reactome pathways enrichment scores (ES) and differential abundance scores (PAS) calculated from transcripts and metabolites log2fold changes
Table S5: GPCR based survival analysis - cancer tissue/TCGA molecular subtypes
Table S6: ligand based survival analysis - cancer tissue/TCGA molecular subtypes
Table S7: enzyme based survival analysis - cancer tissues/TCGA molecular subtypes
Table S8: GPCR-ligand based survival analysis - cancer tissues/TCGA molecular subtypes
Table S9: GPCR-enzyme based survival analysis - cancer tissues/TCGA molecular subtypes
Table S10: GPCR ligands activity from PRISM cancer cell line drug repurposing hub

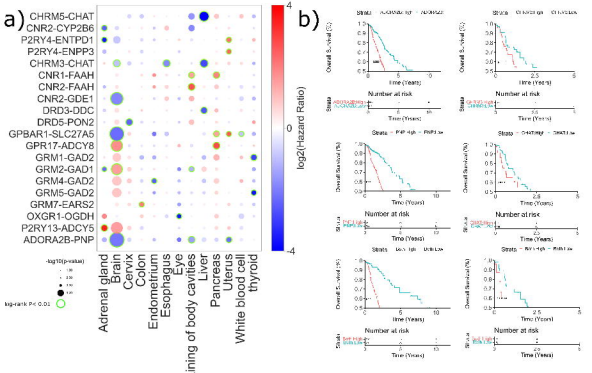

a)
Differential Expression Analysis
TCGA vs. GTEX
Count data from UCSC Xena
TCGA GTEx RNAseq TOIL 19039 samples
16 Primary sites, 76 Subtypes
Pathway/Network Analysis
WebGestalt GSEA
Enrichment Score
Enriched Pathways
Co-Differential Regulation
Receptor
Ligand
Survival Analysis
Patients
Individual Gene
Combination
Receptor Only
Ligand Only
Receptor & Ligand
Advanced Risk Classification
Ligands
GPCR
Transducer
Downstream Signaling
GPCR Centric Network
Signaling Axes Analysis

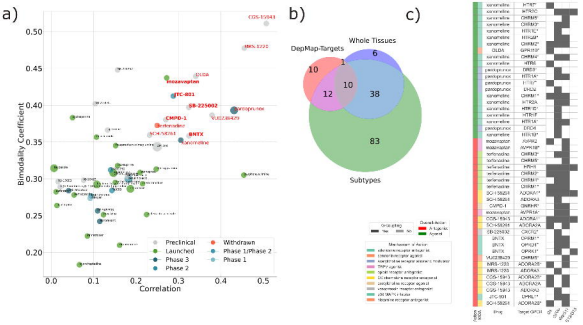

b)
GPCRs
Ligands
Enzymes
Present
Metabolism
c)
Testis
Ovary
Skin
Uterus
Pancreas
Colon
Brain
Breast
Esophagus
Stomach
Lung
thyroid
Prostate
Liver
Kidney
Bladder
GPCRs
Ligands
Enzymes
d)
Kidney
Stomach
Ovary
Testis
Bladder
Skin
Lung
Pancreas
Breast
Colon
Uterus
Brain
Liver
Prostate
Thyroid
Significant Metabolism Pathways
Significant Signaling Pathways
e)
Enrichment Score


a)
Co-Regulation
Cluster 1
Cluster 2
Count
G12/G13
Gq/G11
Gi/Go
Gs
b)
Tissue-wise Similarity
log2(Ligand-wise Similarity)
mean(Co-up - Co-down)
CCR3
CXCR3
CCR5
CXCR2
CCR1
c)
Receptor|Ligand
CXCR3|CCL11
CXCR3|CCL13
CXCR3|CCL19
CXCR3|CCL20
CXCR3|CCL5
CXCR3|CCL7
CXCR3|CXCL10
CXCR3|CXCL11
CXCR3|CXCL12
CXCR3|CXCL9
CALCRL|ADM
CALCRL|ADM2
CALCRL|CALCA
CALCRL|CALCB
RXFP2|INSL3
RXFP2|RLN2
RXFP2|RLN1
RXFP2|RLN3
Bladder
Brain
Colon
Esophagus
Heart
Liver
Lung
Ovary
Pancreas
Prostate
Skin
Stomach
Testis
Uterus
Thyroid


a)
Enrichment Score
2
0
−2
Subtype
Bladder
Brain
Breast
Colon
Esophagus
Kidney
Liver
Lung
Ovary
Prostate
Skin
Stomach
Thyroid
Uterus
b)
Activation of gene expression by SREBF (SREBP)
Purine catabolism
Pyruvate metabolism
Nicotinamide salvaging
Glycosphingolipid metabolism
Pentose phosphate pathway
Metabolism of polyamines
Nicotinate metabolism
Tryptophan catabolism
Synthesis of bile acids and bile salts via 7alpha-hydroxycholesterol
Sphingolipid de novo biosynthesis
Creatine metabolism
Gluconeogenesis
Citric acid cycle (TCA cycle)
Glyoxylate metabolism and glycine degradation
Purine salvage
alpha-linolenic acid (ALA) metabolism
Recycling of bile acids and salts
Cholesterol biosynthesis
Triglyceride catabolism
Glucagon-like Peptide-1 (GLP1) regulates insulin secretion
Synthesis of Prostaglandins (PG) and Thromboxanes (TX)
Acetylcholine regulates insulin secretion
eNOS activation
Glutathione synthesis and recycling
Synthesis of very long-chain fatty acyl-CoAs
Sulfur amino acid metabolism
Arachidonic acid metabolism
Fatty acyl-CoA biosynthesis
PPARA activates gene expression
Regulation of lipid metabolism by PPARalpha
Carnitine metabolism
Glycolysis
1
0
−1
PAAD|Pancreas
BRCA|Breast
PRAD|Prostate
KIRC|Kidney
BLCA|Bladder
c)
Purine catabolism
Enrichment Score
Pathway Abundance
Score
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Breast
Pancreas
Prostate
d)
Ovary
Uterus
Esophagus
Prostate
Pancreas
Lung
Stomach
Testis
Bladder
Colon
Kidney
Breast
Brain
Liver
Skin
Thyroid
Genes
LFC
e)
LFC
5
0
−5


a)
Low Survival
High Survival
Kidney
Breast
Pancreas
Lining of body cavities
Skin
Head and Neck region
White blood cell
Adrenal gland
Eye
Bladder
Cervix
Lung
Liver
Colon
Stomach
Thyroid
Ovary
Esophagus
Uterus
Endometrium
Rectum
Bile duct
Prostate
Paraganglia
Thymus
Lymphatic tissue
Receptors
Ligands
Enzymes
b)
GPCRs
Ligands
Enzymes
Low-HR
High-HR
Nr. of cancer tissues
c)
-log10(p-value)
log-rank P< 0.01
Adrenal gland
Bile duct
Bladder
Breast
Cervix
Colon
Endometrium
Esophagus
Eye
Head and Neck region
Lining of body cavities
Liver
Lymphatic tissue
Prostate
Rectum
Uterus
White blood cell
thyroid
log2(Hazard Ratio)
d)
log2(TPM+0.001)
Median
High
Low
CALCR
CALCB
Both-High
Both-Low
Strata
Overall Survival (%)
Time (Years)
Number at risk
e)
Subtype
Lung_LUAD.3
Skin_SKCM.BRAF_Hotspot_Mutants
Skin_SKCM.Triple_WT
Kidney_KIRC.2
Ovary_OVCA.Mesenchymal
Lung_LUAD.6
Stomach_GI.GS
Skin_SKCM.RAS_Hotspot_Mutants
Esophagus_GI.ESCC
Breast_BRCA.Her2
Breast_BRCA.Basal
Colon_GI.GS
Breast_BRCA.LumB
Liver_LIHC.iCluster:2
log2(Hazard Ratio)
co-up regulated
co-down regulated
Receptor|Ligand


a)
CHRM5-CHAT
CNR2-CYP2B6
P2RY4-ENTPD1
P2RY4-ENPP3
CHRM3-CHAT
CNR1-FAAH
CNR2-FAAH
CNR2-GDE1
DRD3-DDC
DRD5-PON2
GPBAR1-SLC27A5
GPR17-ADCY8
GRM1-GAD2
GRM2-GAD1
GRM4-GAD2
GRM5-GAD2
GRM7-EARS2
OXGR1-OGDH
P2RY13-ADCY5
ADORA2B-PNP
Adrenal gland
Brain
Cervix
Colon
Endometrium
Esophagus
Eye
Lining of body cavities
Liver
Pancreas
Uterus
White blood cell
thyroid
log2(Hazard Ratio)
4
0
-4
-log10(p-value)
log-rank P< 0.01
b)
Overall Survival (%)
Time (Years)
Number at risk
Strata
Overall Survival (%)
Time (Years)
Number at risk
Strata
Overall Survival (%)
Time (Years)
Number at risk
Strata
Overall Survival (%)
Time (Years)
Number at risk
Strata
Overall Survival (%)
Time (Years)
Number at risk
Strata
Overall Survival (%)
Time (Years)
Number at risk


a)
Rimodality Coefficient
Correlation
Preclinical
Launched
Phase 3
Phase 2
Withdrawn
Phase 1/Phase 2
Phase 1
b)
Whole Tissues
DepMap-Targets
Subtypes
c)
Drug
Target GPCR