Functional map of GPCRs, G proteins and TFs relationship.

MSc Bioinformatics – Bruna Plateroti

Table of Contents

Abstract

G-protein coupled receptors (GPCRs) represent the largest family of membrane receptors and are crucial for numerous cellular signalling pathways that regulate vital processes such as cell growth, division, apoptosis and hormonal responses. These receptors, characterised by their seven transmembrane domains, interact with heterotrimeric G-proteins, which transmit signals from extracellular ligands to intracellular effectors, leading to the production of second messengers and activation of the downstream signalling cascades. Small G proteins, functioning as monomeric entities, act as molecular switches within these pathways, playing pivotal roles in cytoskeletal dynamics, vesicle trafficking, and cell proliferation. Transcription factors (TFs), which regulate gene expression by binding to specific DNA sequences, are often dysregulated in cancer, contributing to oncogenesis through aberrant control of genes involved in cell cycle progression, apoptosis evasion, and metastasis.

This study aimed to systematically investigate the association and regulatory relationships between GPCRs, small G-proteins, and TFs in cancer cells using functional pathway inference analysis (FPIA). Leveraging gene dependency data from the Cancer Dependency Map (DepMap) project, the study sought to identify critical genes that exhibit consistent effects on cell viability, providing insights into the molecular mechanisms underlying cancer progression. The research focused on the Ras signalling pathway and serotonin receptors (HTRs) to explore their roles in oncogenesis and metastasis.

Through Pearson's correlation analysis and hierarchical clustering, significant correlations between GPCRs, small G-proteins and TFs were identified, suggesting a complex regulatory network in cancer. Notably, GPCRs such as LPAR1 and LPAR6 were strongly correlated with Ras pathway genes (e.g., NRAS, KRAS), implicating these receptors in Ras-driven oncogenic signalling. Additionally, the study revealed an intriguing association between HTRs and TFs, indicating potential roles for these receptors in cancer cell proliferation, invasion, and angiogenesis. Enrichment analysis highlighted the involvement of miRNA-related pathways in HTR signalling, suggesting novel mechanisms by which HTRs may influence cancer progression.

The findings underscore the importance of understanding GPCR signalling in cancer, with significant implications for the development of targeted therapies. By identifying key molecular interactions, this study paves the way for future research aimed at validating these correlations and exploring the broader landscape of GPCR signalling in cancer. The identification of novel therapeutic targets, particularly within the Ras pathway and serotonin receptor families, could lead to more effective and personalised cancer treatment.

1. Introduction

G-protein coupled receptors (GPCRs) represent the largest family of membrane receptors, with more than 800 members identified in humans (Fredriksson *et al.*, 2003). These receptors are characterised by seven transmembrane domains and are involved in the transduction of extracellular signals into intracellular responses, regulating various physiological processes, including immune responses, metabolism, and sensory perception (*Seven-transmembrane receptors | Nature Reviews Molecular Cell Biology*, no date). As the name implies, these receptors are associated with heterotrimeric G-proteins, which are activated when an extracellular ligand (e.g., hormone, neurotransmitter, growth factor) binds to GPCR. This binding event triggers a conformational change in the receptor that allows them to interact with and activate G-protein (Rosenbaum, Rasmussen and Kobilka, 2009).

Small G-proteins, characterised by their low molecular weight (Mw), are composed of alpha, beta and gamma subunits and act as molecular switches that trigger downstream signalling pathways. These pathways produce second messengers, such as cyclic AMP (cAMP), inositol triphosphate (IP3), and diacylglycerol (DAG), which further propagate the signal within the cell (Wennerberg, Rossman and Der, 2005). These signalling cascades ultimately influence diverse cellular processes, including cell growth, cell division, apoptosis and hormonal response, making GPCRs integral to maintaining cellular homeostasis. Unlike heterotrimeric G-proteins, small G-proteins function as monomeric entities and are involved in regulating a variety of cellular functions, such as cytoskeletal dynamics, vesicle trafficking, and cell proliferation. Small G-proteins, including members of the Ras, Rho, Rab, and Arf families, act as binary molecular switches, cycling between an active GTP-bound state and an inactive GDP-bound state. Their activity is tightly regulated by guanine nucleotide exchange factors (GEFs), which promote the exchange of GDP for GTP, and GTPase-activating proteins (GAPs), which accelerate GTP hydrolysis, returning the protein to its inactive state.

Transcription factors (TFs), on the other hand, are proteins that regulate gene expression by binding to specific DNA sequences, often in the promoter or enhancer regions of target genes. TFs play an essential role in controlling the transcription activity of genes by either facilitating or inhibiting the binding of RNA polymerase to the DNA template. This regulation is essential for cellular differentiation, development, and response to external stimuli. In the context of cancer, TFs are often dysregulated, leading to aberrant expression of genes that promote oncogenesis, such as those involved in cell cycle progression, apoptosis

evasion, and metastasis (Lee and Young, 2013). GPCRs can regulate TF activity through intracellular signalling cascade after ligand binding. Although less direct, TF can regulate the expression of GPCR genes by binding to the promoter regions of GPCR genes and either activating or repressing their transcription. This can occur as part of a feedback loop where GPCRs signalling alters TF activity, which in turn modulates the expression of the GPCR itself or other related receptors.

As noted by (Chaudhary and Kim, 2021), the roles of GPCRs, small G-proteins, and TFs in cancer are particularly significant, as they contribute to cancer initiation, progression, and survival. GPCRs can regulate immune cell recruitment, stromal cell behaviour, and angiogenesis within the tumour microenvironment, creating conditions that promote tumour survival and resistance to therapies (Hanahan and Coussens, 2012), tumour cell invasion, and metastasis through cytoskeletal alterations. Given the vital role of GPCRs in cellular signalling, it is not surprising that they are the most targeted receptor in drug discovery. Nearly 34% of all FDA-approved drugs act on 108 GPCRs (Hauser et al., 2017). This makes understanding their role in cancer signalling critical for identifying novel therapeutic targets and improving existing treatments. GPCRs can function both as oncogenes and tumour suppressors, depending on the cellular context and the specific receptor involved. For instance, the LPAR family has been implicated in various cancers, including breast, ovarian, and lung cancers, where they promote cell proliferation, survival, and metastasis (Mills and Moolenaar, 2003). LPAR1 and LPR6, in particular, have been shown to activate downstream signalling pathways such as the Ras-MAPK and PI3K-AKT pathways, which are crucial for oncogenic transformation (Murph *et al.*, 2007).
Conversely, other GPCRs may act as tumour suppressors by inhibiting proliferative signalling or promoting apoptosis, though these roles are less commonly studied. Rho family GTPases, including RhoA, Rac1, and CDC42, are critical regulators of cytoskeletal dynamics, cell motility, and adhesion. These processes are essential for cancer cell invasion and metastasis, making Rho GTPasa key players in cancer progression (Heasman and Ridley, 2008). Aberrant activation of Rho GTPases has been associated with increased metastatic potential in various cancers similarly to LPAR family (Karlsson *et al.*, 2009). For example, overexpression of Rac1 has been linked to enhanced cell migration and invasion through the reorganization of the actin cytoskeleton and the formation of lamellipodia (*The interdependence of the Rho GTPases and apicobasal cell polarity - PMC*, no date).
In regard to TFs, these can be dysregulated in cancer through various mechanisms, including gene mutations, chromosomal translocations, and epigenetic modifications. For example, the fusion of the ETS family with the

androgen receptor in prostate cancer leads to the activation of oncogenic transcriptional programs that drive tumour growth (Tomlins *et al.*, 2005). Epigenetic modifications, such as DNA methylation and histone acetylation, can also alter TF binding to DNA, resulting in the silencing of tumour suppressor genes or the activation of oncogenes (*Epigenetics in Cancer | New England Journal of Medicine*, no date).

This research aims to systematically investigate and identify associations and regulatory relationships between GPCRs, small G proteins, and TFs in cancer cells using functional pathway inference analysis (FPIA). This approach leverages gene dependency data to assess which genes impact cell viability when targeted across a genome-wide scale, identifying those whose inhibition results in similar survival outcomes across various cell lines. The gene dependency data utilised in this study is derived from the Cancer Dependency Map (DepMap) project. DepMap provides a comprehensive dataset that reveals the dependencies of 700 cancer cell lines, linking gene function to cell survival across diverse tumour types. By integrating DepMap data into the FPIA framework, critical genes can be identified that exhibit consistent effects on cell viability, offering a deeper insight into the molecular mechanisms underlying cancer progression. To assess FPIA, the cordial package (Badshah and Cutillas, 2023) is used, a newly developed package that performs Pearson's correlation analysis of datasets or specific targets. The primary objectives of this study are to identify significant correlations between these molecules across different cancer cell lines, representing 27 tumour lineages; the goal is to uncover statistically significant correlations between these molecules. This process involves determining how the presence or activity of one molecule (e.g. GPCR) correlates with the expression or function of another molecule (e.g., TF or small G-protein). Once the significant correlations are identified, the study will visualise these molecular interactions using heatmaps and network analysis. Heatmaps will provide a more comprehensive view on how specific GPCRs, TFs and small G proteins are interconnected, highlighting clusters or patterns that suggest coordinated activity or functional relationships within cancer cells. Network analysis will provide a more dynamic view, mapping out the complex web of interactions and illustrating how different molecules interact with one another in cancer signalling pathways. The final objective of the study is to interpret how the identified molecular interactions contribute to cancer progression and survival. By understanding the functional relationships between these molecules, the study seeks to shed light on how these molecules may drive tumour growth, metastasis or resistance to therapy. Furthermore, these findings could significantly impact drug development by identifying genes essential for cancer

cell survival, paving the way for the creation of tailored therapies. By uncovering specific weaknesses in various cancer types, this research highlights potential therapeutic targets that could lead to more personalised and effective treatments.

To demonstrate the research approach, the Ras signalling pathway is examined, focusing on the isoforms NRAS, HRAS and KRAS, which are frequently implicated in cancers such as melanoma, lung and colon cancer (Simanshu, Nissley and McCormick, 2017). Ras gene mutations are among the most common oncogenic alterations in human cancers, with KRAS being the most frequently mutated isoform (85%), followed by NRAS (12%) and HRAS (3%)(Cox and Der, 2010). Additionally, MRAS and CDC42 were included in the study for their role in tumorigenesis and metastasis mechanisms. Dysregulation of the Ras pathway can lead to uncontrolled cell proliferation and survival, ultimately driving tumour growth. Despite being one of the most frequently mutated oncogenes in human cancer, directly targeting Ras has proven difficult due to its high affinity for GTP/GDP and the lack of binding pocket, making it “undruggable” (Moore et al., 2020). However, recent discoveries, such as the development of KRAS G12 C inhibitors, offer new hope for targeting Ras-driven cancer (Canon et al., 2019). To further explore these mechanisms, three subsets of five key genes from TFs, G-proteins, and GPCRs were created, each targeting critical components of the Ras signalling network. This multi-gene approach enables a comprehensive investigation of how Ras pathway dysregulation influences oncogenesis through complex gene interactions and signalling crosstalk.

Additionally, the study explores less extensively studied GPCRs, specifically serotonin receptors (HTRs) such as HTR2A, HTR2B, HTR1A and HTR4. While these receptor subtypes involve different signal pathways, they are frequently associated with various cancers. HTR receptors are expressed across different tissues, and the tumour progression is often linked to the dysregulation of serotonin expression patterns (Sarrouilhe et al., 2015). Research has demonstrated that serotonin receptors can stimulate cancer cell proliferation, invasion and tumour angiogenesis (Balakrishna et al., 2021). The identification of key GPCRs, small G proteins, and TFs involved in cancer progression could have significant clinical implications. By mapping these interactions, this study could uncover novel signalling pathways, identify therapeutic targets that modulate receptor activity, and contribute to the development of biomarkers for the early detection of specific cancers. The study also incorporates enrichment analysis to further elucidate the functional significance of serotonin receptors (HTRs) in cancer. Enrichment analysis identifies biological pathways that are

significantly overrepresented among the genes correlated with HTRs, providing insights into the molecular mechanisms by which these receptors influence cancer progression.

2. Materials and Methods

2.1 Data availability

The data utilised in this study is derived from the Cancer Dependency Map (DepMap) project, a collaboration between the Broad Institute (Boston, USA) and the Wellcome Sanger Institute (UK) (The Cancer Dependency Map at Sanger - The Cancer Dependency Map at Sanger, no date). The Cancer Dependency Map aims to identify and catalogue genetic dependencies in cancer cells across various types of cancers, providing a rich resource for understanding cancer biology and identifying potential therapeutic targets. The specific dataset used, DepMap RNAi 2019, includes comprehensive genetic information for 706 cell lines, represented in rows, and 17,313 genes, represented in columns. Each cell line is uniquely identified by a DepMap_ID and associated metadata. A separate data set (Selected genesets) consisting of two columns, one for gene sets (e.g. RTKs, GPCRs, G-proteins, TFs, etc.) and another for genes with the list of genes part of that set.

2.2 Data manipulation

Data manipulation was performed using R programming with the RStudio (v. 1.4.9.4) environment. Genes for GPCRs, TFs and small G proteins were selected from the Selected genesets data and cross-referenced with the DepMap RNAi dataset. To focus on genes with a meaningful impact on cell viability, only those with a DepMap score equal to or higher than 0.5 in at least 10% of cell lines were retained. This threshold was chosen based on previous literature, such as the study by (Tsherniak et al., 2017) which demonstrated that this cutoff effectively identifies essential genes in cancer cells.

2.3 Correlation analysis

Functional pathway inference analysis (FPIA) was employed to explore the functional relationships between genes by analysing their co-expression patterns across 706 cancer cell lines. The analysis was conducted using Pearson's correlation, a widely used statistical measure that quantifies the strength of association between two continuous variables. In this study, Pearson's correlation was applied in parallel across the dataset, allowing for a comprehensive analysis of gene-gene interactions. The approach was designed

to identify potential pathways that are co-regulated in cancer, providing insights into the underlying molecular mechanism of cancer progression. This method has been applied in recent studies, such as those by (Badshah and Cutillas, 2023)), to uncover novel gene networks involved in various biological processes. The correlation analysis was performed using the 'cordial' package, specifically developed for this purpose and obtained from the CutillasLab GitHub repository (CutillasLab, 2024). The package provides several functions, including 'cor_map()', 'cor_target_map()' and 'cor_targets()', which facilitate the calculation of pairwise correlations and the identification of key gene interactions.

The filtered dataset, containing only genes with significant impact in at least 10% of cancer cells, was used to generate three data frames: one with all combined data, another with GPCRs vs. G-protein, and a third with GPCRs vs. TFs. To enhance the efficiency of the analysis, these data frames were converted into data table format using the 'data.table' package, which offers faster data manipulation capabilities compared to traditional data frame . The 'cor_map()' function was employed to calculate pairwise correlation for all targets in the dataset (GPCRs vs G-protein vs TFs), providing a comprehensive map of gene interactions. Simultaneously, 'cor_target_map()' was used to focus on specific gene pairs, such as GPCRs vs. G-protein and GPCRs vs. TFs, allowing for a more targeted analysis. The resulting output consisted of three matrices, each containing the target genes and their correlated partners, along with the corresponding r-values, p-values, and q-values (excluding self-correlation). These matrices were saved as .csv files in the designated data output folder for further analysis.

2.4 Hierarchical cluster analysis

Hierarchical cluster analysis was employed to visualise the complex relationships between the different groups of genes, specifically GPCRs vs. G-proteins and GPCRs vs. TFs. This method provides a powerful visual representation on how genes cluster based on their correlation patterns, revealing potential functional relationships that might not be evident from correlations values alone. The analysis was performed using 'ComplexHeatmap' package from Bioconductor, which is well-suited for creating detailed and customised heatmaps. Before generating the heatmaps, the data was rigorously filtered to include only significant correlations ($q < 0.05$), thereby reducing the risk of over-clustering and ensuring that visualisation focuses on the most relevant gene interactions. The data was converted from wide to long format using the 'dcast()' function from the 'data.table' package and transformed into a matrix format suitable for the 'Heatmap()' function. Thuis transformation was necessary to ensure that the

data was properly structured for hierarchical clustering, which relies on matrix input.

2.5 NRAS has an example

To demonstrate the methodology used in this study, the isoforms NRAS, KRAS, and HRAS were selected, key subtypes of G-proteins involved in the Ras signalling pathway. This pathway is critical in regulating cell growth and differentiation, and mutations in these genes are commonly associated with various cancers. The study used network information from the KEGG (Kyoto Encyclopaedia of Genes and Genomes) website to highlight the applicability of this approach. Three subsets of genes were created, each containing five key genes relevant to TFs (PML, RUNX1, TP53, ATF1, ADCYAP1R1), G-proteins (HRAS, NRAS, KRAS, MRAS, CDC42) and GPCRs (LPAR1, LPAR2, LPAR5, LPAR6, S1PR1). Following the data manipulation and correlation analysis procedures described earlier, these subsets were further filtered to retain only those gene pairs with q-values < 0.05. The top ten positive and negative correlations were selected from the result matrix, providing a focused view of the most significant interactions. These values were combined into a single variable named 'results', which included an additional column called 'Direction' to indicate whether the correlation was positive or negative. A custom function was developed to generate a bar plot of these results using 'ggplot2' package, which is known for its versatility in creating publication-quality visualisations. The plot was saved as a .pdf file, providing a visual summary of the key findings from the study.

2.6 Functions design

To streamline the analysis process, three custom functions were developed for this study: 'find_the_g_protein ()', 'find_the_tf ()' and 'find_the_gpcrs()'. Each function was designed to facilitate specific tasks within the analysis. 'Find_the_g_protein ()' takes GPCRs as input and returns the associated G-proteins as a network diagram, which is useful for visualising the interactions between these genes. 'Find_the_tf ()' takes GPCRs and returns the associated TFs in a summary table, providing a quick reference for researchers interested in transcriptional regulation. 'Find_the_gpcrs()' takes TFs as input and returns the associated GPCRs in a summary table, offering insights into the potential downstream effects of TFs on GPCRs activity. Each function requires four arguments: the gene of interest (parameters 'gpcr_gene' and 'tf_gene'), an input dataset (parameter 'matrix'), a dataset containing only the genes of interest (parameters 'gp_gene', 'tf_gene', 'gpcr_gene') and a threshold of 0.05 to ensure that only significant correlations are considered.

2.7 HTR receptors as insight

To further apply our approach to an interesting set of genes, the serotonin receptors (HTRs) were selected for a focused analysis. Specifically, HTR2A, HTR2B, HTR1A, HTR1B, HTR1D, and HTR4 were chosen to better understand the overview of the role of serotonin in different aspects of cancer biology and their different role in various cancers. These receptors were extracted from the original data, and only those with significant impact on cancer cell lines were retained (HTR2A, HTR2B, HTR1A, HTR4). These selected HTR genes were combined with the G-proteins and TFs previously identified during data manipulation to create a comprehensive dataset for further analysis. This dataset was then converted into a data table for FPIA using the 'cordial' package. The correlation analysis was performed on this combined dataset, and the results were saved as a .csv file for further examination. A heatmap was generated using the 'ComplexHeatmap' package, which visually depicted the relationship between HTRs, G-proteins and TFs. Additionally, a network graph was produced using only the first twenty positive correlations ($q < 0.05$), providing a clear visualisation of the most significant interactions. The analysis offered additional insights into the potential roles of HTRs in cancer biology, highlighting their interaction with other gene categories in the context of cancer cell signalling.

Network graph analysis was conducted using the 'igraph' and 'dplyr' packages in R to explore the interactions between HTRs, G-proteins and TFs. The correlation data was filtered to include only§ significant correlations ($q<0.05$). To focus specifically on the connections of individual receptors, the data was further filtered to isolate correlations where HTR4 or HTR2A was the target gene. This approach allowed for the creation of detailed network graphs that highlight the direct interactions of these receptors, providing insights into their potential roles within the broader signalling networks involved in cancer-related processes. Finally, an enrichment analysis was performed using the 'ClusterProfiler' package in R to identify which biological pathways are significantly enriched among the correlated gene sets. The 'ClusterProfiler' package facilitate the statistical analysis and visualisation of functional profiles for genes and gene clusters, providing insights into the biological processes, molecular functions, and pathways that are overrepresented in the dataset. The enrichment analysis results were visualised using dot plots, which were saved as a .pdf file for inclusion in the study's supplementary materials.

3. Results

3.1 Overview of global interactions

To effectively illustrate the global interactions between GPCRs, small G-proteins and TFs in cancer cells, a clustered heatmap was generated. The heatmap includes only GPCRs, G-proteins and TFs that exhibit a significant impact on cancer cell viability, as determined by their DepMap scores (Fig.1 and Fig.2).

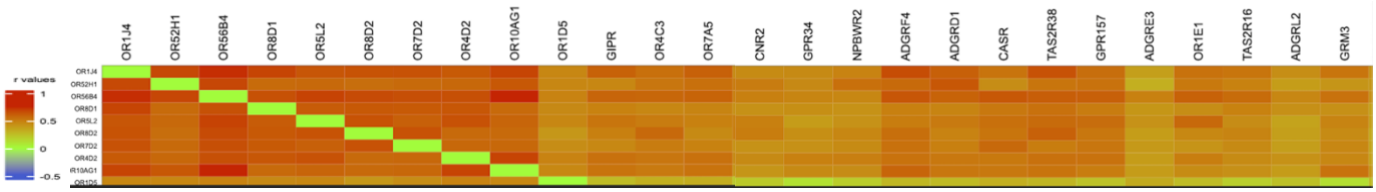

Figure n.1, zoomed in section of the clustered heatmap of G-protein coupled receptors (GPCRs) on rows vs. small G proteins in columns, showing the strongest correlation between these proteins. The heatmap uses a colour gradient ranging from red to bright green to represent the intensity of the Pearson's correlation coefficient, with red indicating strong positive correlations (around 1) and green indicating negative correlations (down to -0.5). Bright green regions correspond to missing values (N/A) imputed as zero during heatmap generation. Given the extensive size of the dataset, only the most relevant correlations are shown, enabling the identification of specific patterns and relationships.

As shown in both heatmaps, there are distinct clusters of gene interactions where GPCRs exhibit strong correlations with G-proteins (Fig.1) and TFs (Fig.2).

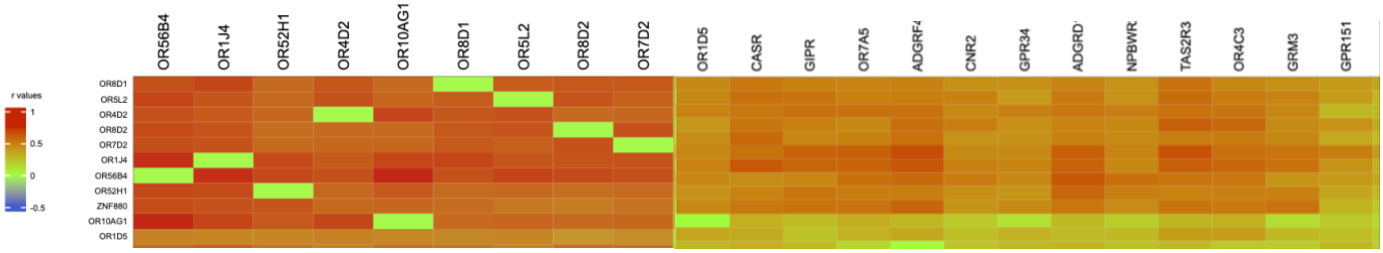

Figure n.2, zoomed in section of the clustered heatmap of G-protein coupled receptor (GPCRs) on rows vs. transcription factors (TFs) on columns, showing the strongest correlation between these proteins. The heatmap uses a colour gradient ranging from red to bright green to represent the intensity of the Pearson's correlation coefficient, with red indicating strong positive correlations (around 1) and green indicating negative correlations (down to -0.5). Bright green regions correspond to missing values (N/A) imputed as zero during heatmap generation. Given the extensive size of the dataset, only the most relevant correlations are shown, enabling the identification of specific patterns and relationships.

For instance, clusters of highly correlated pairs (depicted in red) suggest functional relationships between these molecules. Notably, GPCRs such as OR10AG1 and OR1J4 positively correlated with small G-proteins like OR56B4, indicating potential regulatory mechanisms that may contribute to signal transduction. Conversely, non-significant correlations showed in lighter green reflect isolated interactions or weak associations, which may still be biologically relevant in specific contexts.

3.2 Detailed analysis of critical pathways

To further explore these global interactions, a focused analysis was conducted on the Ras signalling pathway, one of the most frequently altered pathways in cancer. This pathway is often dysregulated due to mutations in key genes such as KRAS, NRAS, and HRAS. A bar plot was generated to visualise the correlations between Ras pathway genes and selected GPCRs and TFs (Fig.3). As shown in Figure 3, five key genes from GPCR and TF subset were chosen to provide a more targeted exploration of these interactions. The analysis reveals distinct patterns of interaction between Ras pathway genes and GPCRs.

For instance, as shown in Figure 3, LPAR1 and LPAR6 show a positive correlation with NRAS, suggesting a potential regulatory role in the Ras signalling pathway. These correlations may indicate that LPAR family receptors could influence downstream signalling, potentially impacting tumour growth or survival. Additionally, TFs such as ADCYAP1R1 displayed correlations with MRAS and NRAS, further implicating transcriptional regulation as a key modulator in the Ras-driven oncogenic process.

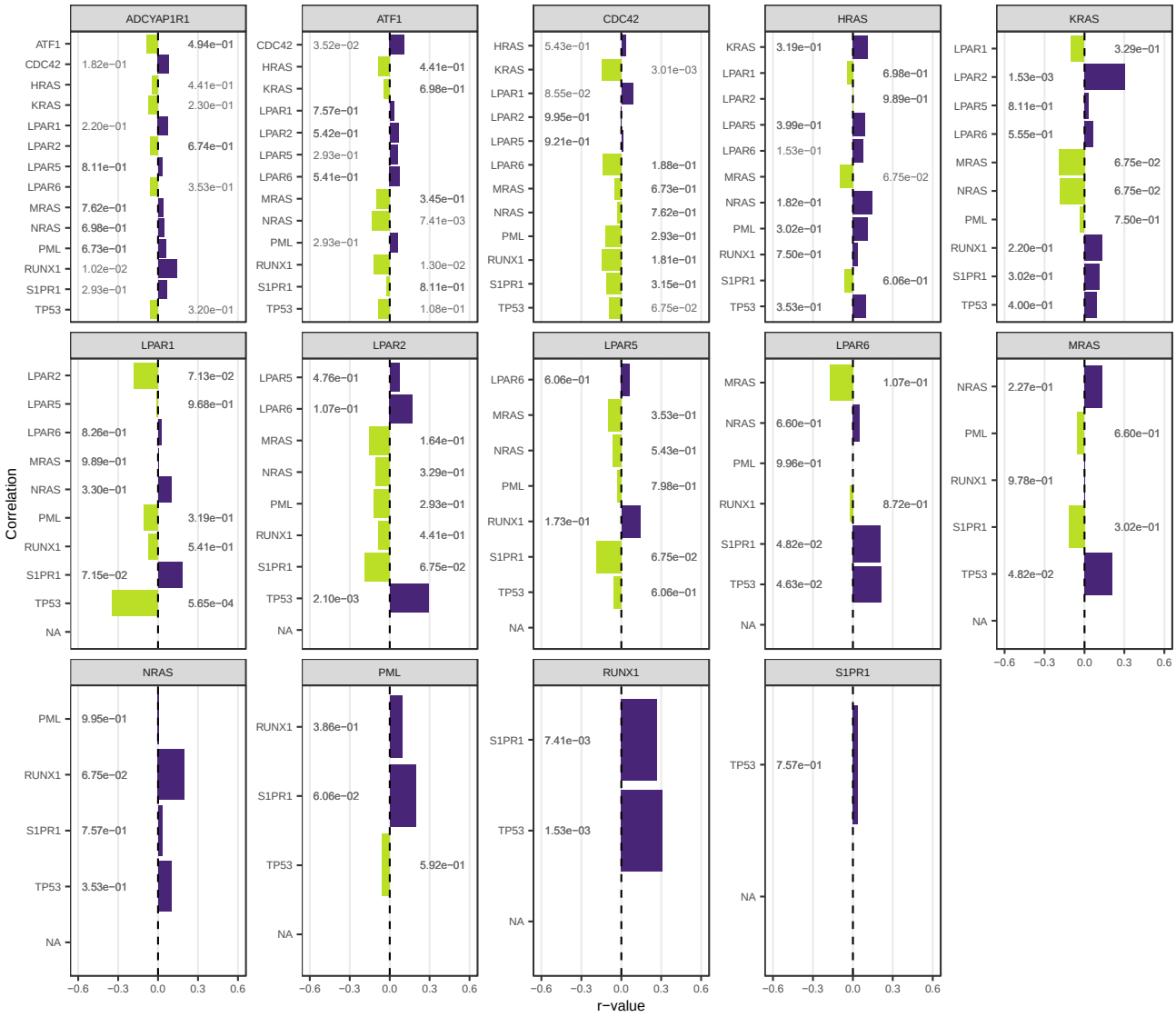

Figure n.3, illustrate a bar plot showcasing the strongest positive and negative correlations (top positive and bottom negative) identified within three key gene subsets. The analysis focuses on genes involved in the Ras signalling pathway, including NRAS, KRAS, HRAS, MRAS and CDC42, in conjunction with selected GPCRs (LPAR1, LPAR2, LPAR5, LPAR6, S1PR1) and TFs (PML, RUNX1, TP53, ATF1, ADCYAP1R1). The X-axis represents the r-values, which quantify the strength and direction of the correlations, blue bars positive and green bars negative. The Y-axis lists the corresponding genes involved in these correlations.

3.3 Case study of serotonin receptors (HTRs)

A case study was performed on the serotonin receptor (HTR) family to explore their interaction with G-proteins and TFs in cancer cells. HTR receptors are known to influence various cellular processes, including proliferation, invasion, and angiogenesis, making them relevant targets for cancer research. The heatmap shown in Figure 4, reveals intriguing clusters of strong correlation between HTR receptors and specific TFs.

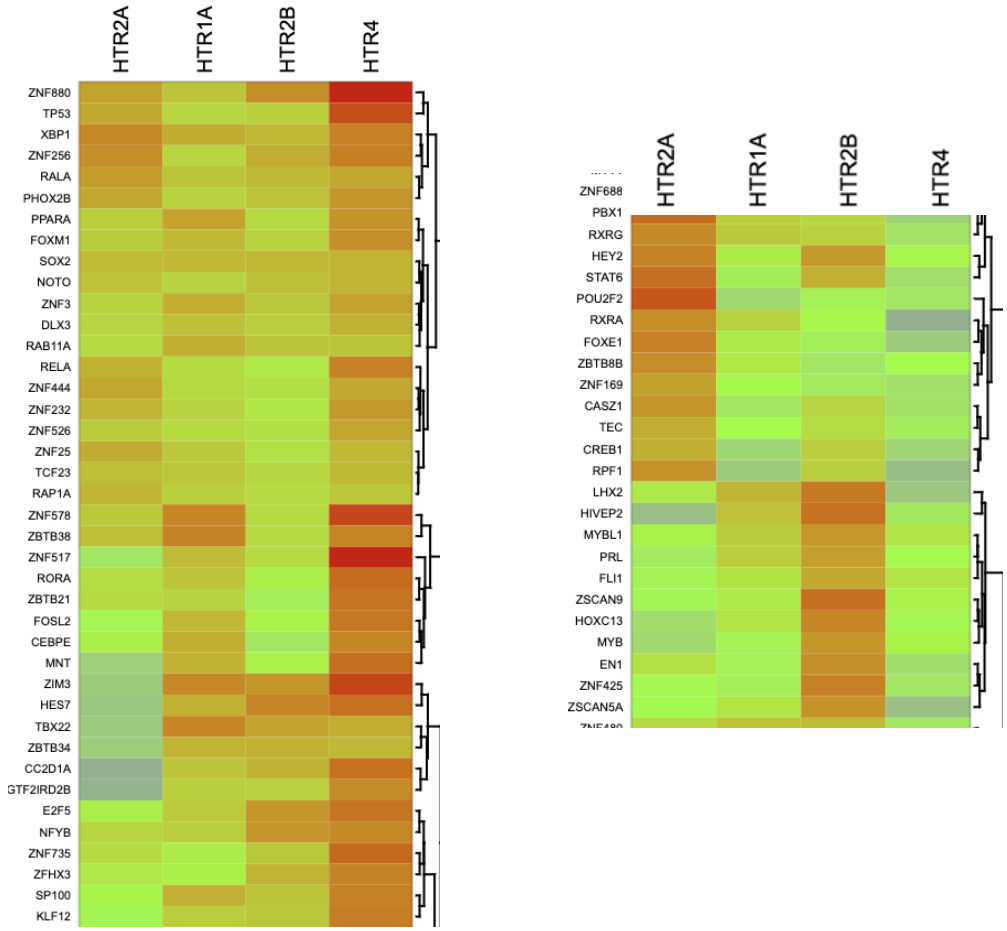

Figure n.4, zoomed in section of the clustered heatmap, focusing on the key interactions between four serotonin receptors (HTR: HTR2A, HTR2B, HTR1A, and HTR4) and their associated transcription factors (TFs) and G-proteins. The heatmap uses a colour gradient ranging from red to bright green to represent the intensity of the Pearson's correlation coefficient, with red indicating strong positive correlations (around 1) and green indicating negative correlations (down to -0.5). Bright green regions correspond to missing values (N/A) imputed as zero during heatmap generation. Given the extensive size of the dataset, only the most relevant correlations are shown, enabling a clearer interpretation of the key relationships.

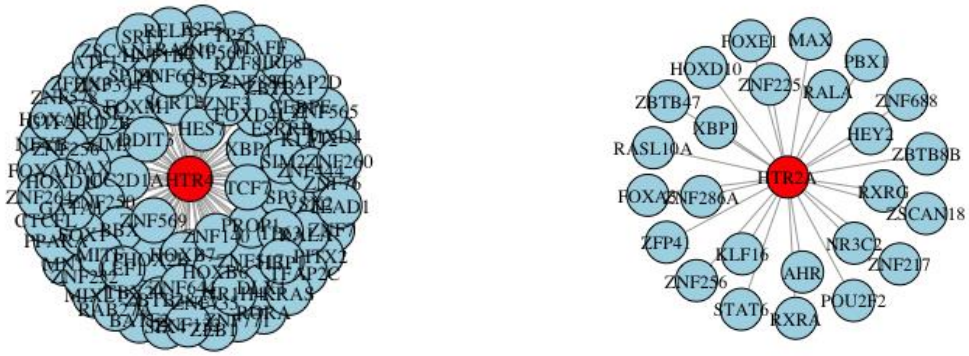

Figure n.5 illustrates the network graphs for the receptors HTR4 (left) and HTR2A (right), showcasing only direct connection to each primary node. The data used for these visualisations were filtered to include only significant correlations (q <0.05).

Network graphs were generated to complement the heatmap visualisation, offering an alternative perspective on the interactions among receptors, G-protein, and TFs. As illustrated in Figure 5, HTR4 has significantly more interactions with G-proteins and TFs compared to HTR2A. This suggests that HTR4 may play a more central role in the signalling network associated with these pathways, potentially indicating a broader influence on cancer-related processes.

3.4 Enrichment analysis of serotonin receptors (HTRs)

The enrichment analysis shown in Figure 6 revealed that the pathways with the highest gene ratios were the pattern specification process and cell fate commitment. These pathways are critical in development biology and are often implicated in cancer progression, as they involve the process by which cells are directed to adopt specific fates, a key aspect of tumorigenesis. The enrichment analysis of these pathways suggests that HTRs may play significant roles in influencing the developmental-like process that occur in cancer cells, contributing to the aberrant cell fate decisions that drive cancer progression.

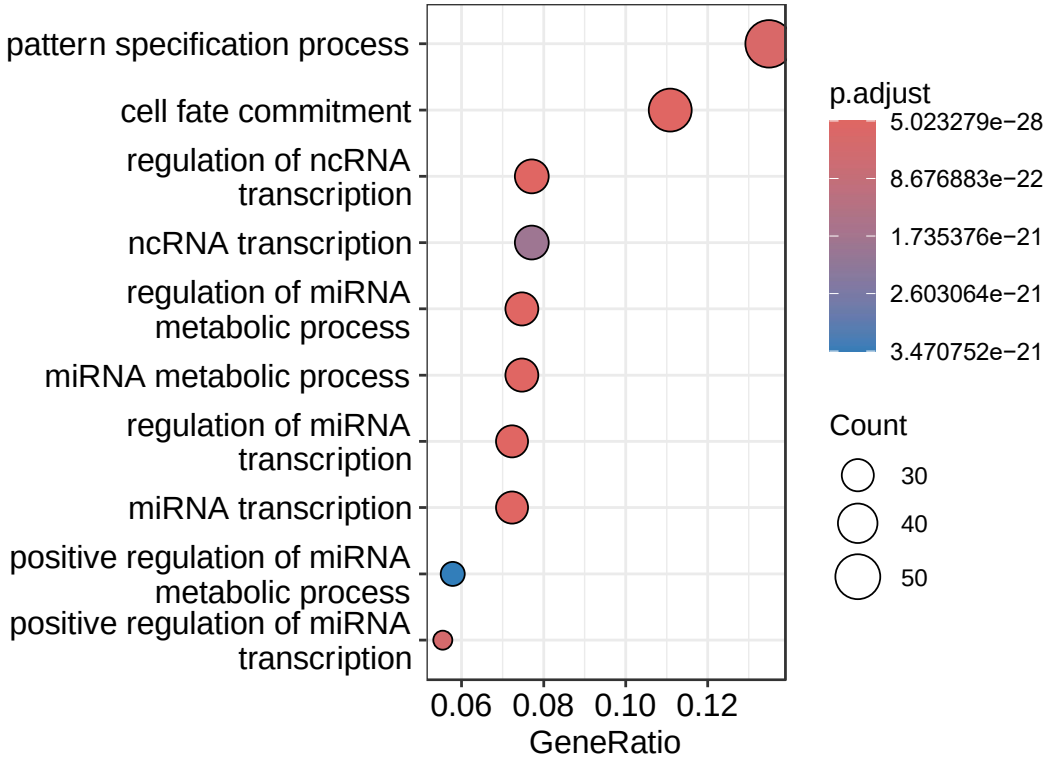

Figure n.6 represents the results of a Gene Ontology (GO) term enrichment analysis conducted on a gene set comprising HTRs (HTR2A, HTR2B, HTR1A, HTR4), G-protein, and TFs using the R package "clusterProfiler". The X-axis represent the Gene Ratio indicating the proportion of the genes in the analysed set that are associated with specific GO term. The Y-axis lists the specific GO biological processes that were found to be significantly enriched. The colour gradient of the dots corresponds to the adjusted p-values, with red indicating strong enrichment and blue indicating weaker or non-significant enrichment. The size of the dot reflects the number of genes associated with GO term.

Additionally, other pathways with lower gene ratios but still highly significant p-values included regulation of miRNA metabolic process, miRNA metabolic process, regulation of miRNA transcription, and miRNA transcription. miRNAs (microRNAs) are small non-coding RNA that regulate gene expression during post-transcription processes and are highly recognised for their role in cancer biology, particularly in the regulation of oncogenes and tumour suppressor genes. The enrichment of these miRNA-related pathways suggests that HTRs may be involved in gene expression through miRNA regulation, potentially influencing cancer cell behaviour.

4. Discussion

4.1 Overview of findings

In this study, the functional relationship between G-protein-coupled receptors (GPCRs), small G-protein, and transcription factors (TFs) in cancer cells were explored using functional pathway inference analysis (FPIA). By leveraging the comprehensive dataset from the Cancer Dependency Map (DepMap) project, significant correlations were identified, suggesting intricate regulatory networks that drive cancer progression. The findings emphasised the crucial role of GPCRs in modulating the activity of small G-proteins and TFs and their involvement in vital signalling pathways such as the Ras pathway. Additionally, the focused analysis on serotonin receptors (HTRs) highlights their potential role in tumour progression through interactions with TFs and small G-proteins, with implications for cell fate determination and miRNA regulation.

4.2 GPCRs and cancer signalling networks

GPCRs are known to play a key role in cancer biology due to their extensive involvement in cell signalling pathways that regulate proliferation, apoptosis and migration. The current study's findings validate these roles, particularly in the context of the Ras signalling pathway, which is often dysregulated in various cancers (Rosenbaum, Rasmussen and Kobilka, 2009). Strong correlations were observed between GPCRs such as LPAR1 and LPAR6 and key Ras pathway genes, including NRAS and KRAS. These findings are consistent with previous studies that implicated the LPA (lysophosphatidic acid) signalling axis in cancer cell survival and metastasis (Mills and Moolenaar, 2003). It is suggested that these LPARs could influence Ras-driven oncogenic processes, making them potential targets for therapeutic intervention, especially in cancers with Ras mutations.

4.3 Role of small G-protein in cancer

Small G-proteins act as molecular switches in signalling pathways, controlling various cellular processes essential for cancer development. Significant correlations were identified between small G-proteins like MRAS and CDC42 with both GPCRs and TFs, showing their central role in modulating these signalling networks. CDC42, for instance, has been extensively documented for its role in regulating the cytoskeleton, cell polarity, and migration processes that are crucial for cancer cell invasion and metastasis (Heasman and Ridley, 2008).
The analysis revealed that MRAS exhibit a positive correlation with the tumour suppressor TP53, suggesting that MRAS may interact with TP53-related pathways, potentially influencing cell cycle regulation and apoptosis. This interaction may reflect a regulatory mechanism where MRAS contributes to

TP53-mediated tumour suppression or, alternatively, where TP53 modulates MRAS activity as part of a broader cellular response to stress or DNA damage (The Old and the New in p53 Functional Regulation - ScienceDirect, no date). Additionally, CDC42 was found to have a positive correlation with the GPCR LPAR1. This correlation indicates a potential link between CDC42 and LPAR1 in driving processes like cell migration and invasion, which are critical for cancer metastasis. LPAR1 is known to be involved in the lysophosphatidic acid (LPA) signalling pathway, which plays a key role in promoting cell survival, proliferation and motility. The positive correlation between CDC42 and LPAR1 suggests that the activation of CDC42 may be influenced by LPAR1 signalling, thereby contributing to the aggressive behaviour of cancer cells. The interplay between GPCRs and small G-proteins is suggested to be critical in driving these processes, potentially offering new avenues for therapeutic targeting.

4.4 Transcription factors and gene regulation in cancer

Transcription factors are central to gene regulation, and their interaction with GPCRs and small G-proteins appears to form a complex regulatory network that influences cancer progression. Significant correlations were observed between TFs like TP53, POU2F2, and RUNX1 with GPCRs and small G-proteins, highlighting potential feedback loops where GPCR signalling modulates TF activity, which in turn regulates the expression of GPCRs and other critical genes. TP53, a well-known tumour suppressor, showed strong correlations with HTRs, suggesting that these interactions could play a role in tumour suppression mechanisms or, conversely, in the evasion of these mechanisms by cancer cells. Additionally, the correlation between POU2F2 and HTRs suggests that this TF could be involved in the regulation of immune responses within the tumour microenvironment, as suggested by prior studies (Imamdin and van der Vorst, 2023)

4.5 Insights from serotonin receptors (HTRs) analysis

The focused analysis on HTRs provided intriguing insights into their role in cancer. Serotonin receptors, though primarily known for their function in the nervous system, have been increasingly recognised for their involvement in cancer progression (Sarrouilhe et al., 2015). Strong correlations between HTR2A, HTR2B, and HTR4 with specific TFs and G-proteins were observed, suggesting that these receptors may contribute to cancer cell proliferation, invasion, and angiogenesis. The enrichment analysis further supported this by identifying significant pathways such as pattern specification and cell fate commitment, which are critical in developmental processes and are often co-opted in cancer (Hanahan and Weinberg, 2011).

Although with lower gene rations, the involvement of miRNA-related pathways adds another layer of complexity to the role of HTRs in cancer. miRNA are known

to regulate a wide range of oncogenes and tumour suppressors, and their dysregulation is a hallmark of cancer (Oncomirs — microRNAs with a role in cancer | Nature Reviews Cancer, no date). The correlation observed suggests that HTRs may influence these pathways, potentially through the regulation of miRNA transcription and processing, which could affect the stability and expression of target genes involved in cancer.

4.6 Clinical implication and future directions

Several important clinical implications have emerged from the findings of this study. First, the identification of key GPCRs, small G-proteins, and TFs that interact in cancer cells provides potential targets for therapeutic intervention. Given their accessibility on the cell surface and their role in signalling, GPCRs remain particularly attractive targets for drug development. The correlations observed between Ras pathway genes and GPCRs such as LPAR1 and LPAR6 suggest that these receptors could be targeted to modulate Ras-driven oncogenic signalling, a critical need given the challenges in directly targeting Ras proteins (Moore et al., 2020).

Additionally, the involvement of HTRs in cancer cell signalling and their potential regulation of miRNA pathways highlights new opportunities for targeting these receptors in cancers where serotonin signalling is dysregulated. Given the broad expression of HTRs across tissues, drugs that modulate their activity could have widespread therapeutic applications, particularly in combination with existing treatments to enhance efficacy or overcome resistance.

Future studies should focus on validating these findings in vivo and exploring the mechanistic basis of the observed correlations. Additionally, expanding the analysis to include other GPCR families and their interactions with different signalling molecules could further elucidate the complex network driving cancer progression.

Valuable insights into the functional relationships between GPCRs, small G-proteins and TFs in cancer have been provided by this study. By integrating data from the DepMap project with advanced correlation analysis techniques, key interactions that may drive tumour progression have been identified, and potential targets for therapeutic intervention have been highlighted. The importance of understanding GPCRs signalling in cancer has been underscored, paving the way for future research to develop more effective, targeted treatments.

5. Conclusions

In this study, the intricate interplay between G-protein coupled receptors (GPCRs), small G-proteins, and transcription factors (TFs) in cancer cells was systematically investigated using functional pathway inference analysis (FPIA) combined with comprehensive gene dependency data from the Cancer Dependency Map (DepMap) project. The overarching objective was to uncover significant correlations and regulatory relationships among these molecules to enhance the understanding of their role in cancer progression and survival. The results have demonstrated that the study successfully met the aims by identifying critical molecular interactions and highlighting potential therapeutic targets. However, the research also opens new avenues for future exploration, particularly in the context of validating these findings and extending the analysis to other related signalling networks.

5.2 Future directions and limitations

While this study has made significant strides in understanding the relationships between GPCRs, small G-protein, and TFs in cancer, it also raises several important questions that warrant further investigation. One of the key limitations of this study is the reliance on correlation analysis, which, while powerful, does not establish causality. Future work should focus on experimental validation of the identified correlations to determine whether these interactions are functionally significant in the context of cancer progression. In vivo, studies using cancer models could provide valuable insights into the mechanistic basis of these interactions and help confirm whether the observed correlations translate into functional outcomes in a physiological setting.

Additionally, the study primarily focused on a subset of genes within the Ras signalling pathway and serotonin receptors, which, while relevant, represent only a fraction of the broader landscape of GPCR signalling in cancer. Expanding the analysis to include other GPCRs families (e.g., CXCR4, LPAR1) and their interactions with different signalling molecules could further elucidate the complex network driving cancer progression.

Future research could explore the specific mechanisms by which HTRs influence miRNA activity and how this affects cancer behaviour. This could lead to the identification of novel biomarkers for cancer diagnosis or prognosis and the development of miRNA-targeted therapies.

The study suggested that GPCR signalling might modulate TF activity, which in turn could regulate the expression of GPCRs and other genes involved in cancer. Investigating these feedback mechanisms could provide a deeper understanding

of how GPCR signalling is regulated in cancer and how it contributes to tumour growth, metastasis, and resistance to therapy. Thus, the correlations observed between Ras pathway genes and GPCRs such as LPAR1 and LPAR6 suggest that these receptors could be targeted to modulate Ras-driven oncogenic signalling, a critical need given the challenges in directly targeting Ras protein.

6. References

Badshah, I.I. and Cutillas, P.R. (2023) 'Systematic identification of biochemical networks in cancer cells by functional pathway inference analysis', *Bioinformatics*, 39(1), p. btac769. Available at: https://doi.org/10.1093/bioinformatics/btac769.

Balakrishna, P. *et al.* (2021) 'Serotonin Pathway in Cancer', *International Journal of Molecular Sciences*, 22(3), p. 1268. Available at: https://doi.org/10.3390/ijms22031268.

Canon, J. *et al.* (2019) 'The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity', *Nature*, 575(7781), pp. 217–223. Available at: https://doi.org/10.1038/s41586-019-1694-1.

Chaudhary, P.K. and Kim, S. (2021) 'An Insight into GPCR and G-Proteins as Cancer Drivers', *Cells*, 10(12), p. 3288. Available at: https://doi.org/10.3390/cells10123288.

Cox, A.D. and Der, C.J. (2010) 'Ras history', *Small GTPases*, 1(1), pp. 2–27. Available at: https://doi.org/10.4161/sgtp.1.1.12178.

CutillasLab (2024) 'CutillasLab/cordial'. Available at: https://github.com/CutillasLab/cordial (Accessed: 2 August 2024).

Epigenetics in Cancer | New England Journal of Medicine (no date). Available at: https://www.nejm.org/doi/10.1056/NEJMra072067?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed (Accessed: 23 August 2024).

Fredriksson, R. *et al.* (2003) 'The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints', *Molecular Pharmacology*, 63(6), pp. 1256–1272. Available at: https://doi.org/10.1124/mol.63.6.1256.

Hanahan, D. and Coussens, L.M. (2012) 'Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment', *Cancer Cell*, 21(3), pp. 309–322. Available at: https://doi.org/10.1016/j.ccr.2012.02.022.

Hanahan, D. and Weinberg, R.A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*, 144(5), pp. 646–674. Available at: https://doi.org/10.1016/j.cell.2011.02.013.

Hauser, A.S. *et al.* (2017) ‘Trends in GPCR drug discovery: new agents, targets and indications’, *Nature Reviews Drug Discovery*, 16(12), pp. 829–842. Available at: https://doi.org/10.1038/nrd.2017.178.

Heasman, S.J. and Ridley, A.J. (2008) ‘Mammalian Rho GTPases: new insights into their functions from in vivo studies’, *Nature Reviews. Molecular Cell Biology*, 9(9), pp. 690–701. Available at: https://doi.org/10.1038/nrm2476.

Imamdin, A. and van der Vorst, E.P.C. (2023) ‘Exploring the Role of Serotonin as an Immune Modulatory Component in Cardiovascular Diseases’, *International Journal of Molecular Sciences*, 24(2), p. 1549. Available at: https://doi.org/10.3390/ijms24021549.

Karlsson, R. *et al.* (2009) ‘Rho GTPase function in tumorigenesis’, *Biochimica Et Biophysica Acta*, 1796(2), pp. 91–98. Available at: https://doi.org/10.1016/j.bbcan.2009.03.003.

Lee, T.I. and Young, R.A. (2013) ‘Transcriptional Regulation and its Misregulation in Disease’, *Cell*, 152(6), pp. 1237–1251. Available at: https://doi.org/10.1016/j.cell.2013.02.014.

Mills, G.B. and Moolenaar, W.H. (2003) ‘The emerging role of lysophosphatidic acid in cancer’, *Nature Reviews. Cancer*, 3(8), pp. 582–591. Available at: https://doi.org/10.1038/nrc1143.

Moore, A.R. *et al.* (2020) ‘RAS-targeted therapies: is the undruggable drugged?’, *Nature Reviews. Drug Discovery*, 19(8), pp. 533–552. Available at: https://doi.org/10.1038/s41573-020-0068-6.

Murph, M. *et al.* (2007) ‘Liquid chromatography mass spectrometry for quantifying plasma lysophospholipids: potential biomarkers for cancer diagnosis’, *Methods in Enzymology*, 433, pp. 1–25. Available at: https://doi.org/10.1016/S0076-6879(07)33001-2.

Oncomirs — microRNAs with a role in cancer | Nature Reviews Cancer (no date). Available at: https://www.nature.com/articles/nrc1840 (Accessed: 21 August 2024).

Rosenbaum, D.M., Rasmussen, S.G.F. and Kobilka, B.K. (2009) ‘The structure and function of G-protein-coupled receptors’, *Nature*, 459(7245), pp. 356–363. Available at: https://doi.org/10.1038/nature08144.

Sarrouilhe, D. *et al.* (2015) ‘Serotonin and Cancer: What Is the Link?’, *Current Molecular Medicine*, 15(1), pp. 62–77.

Seven-transmembrane receptors | Nature Reviews Molecular Cell Biology (no date). Available at: https://www.nature.com/articles/nrm908 (Accessed: 23 August 2024).

Simanshu, D.K., Nissley, D.V. and McCormick, F. (2017) ‘RAS Proteins and Their Regulators in Human Disease’, *Cell*, 170(1), pp. 17–33. Available at: https://doi.org/10.1016/j.cell.2017.06.009.

The Cancer Dependency Map at Sanger - The Cancer Dependency Map at Sanger (no date). Available at: https://depmap.sanger.ac.uk/# (Accessed: 21 July 2024).

The interdependence of the Rho GTPases and apicobasal cell polarity - PMC (no date). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4601375/ (Accessed: 23 August 2024).

The Old and the New in p53 Functional Regulation - ScienceDirect (no date). Available at: https://www.sciencedirect.com/science/article/pii/S1077315097926168?via%3Dihub (Accessed: 21 August 2024).

Tomlins, S.A. *et al.* (2005) 'Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer', *Science (New York, N.Y.)*, 310(5748), pp. 644–648. Available at: https://doi.org/10.1126/science.1117679.

Tsherniak, A. *et al.* (2017) 'Defining a Cancer Dependency Map', *Cell*, 170(3), pp. 564-576.e16. Available at: https://doi.org/10.1016/j.cell.2017.06.010.

Wennerberg, K., Rossman, K.L. and Der, C.J. (2005) 'The Ras superfamily at a glance', *Journal of Cell Science*, 118(5), pp. 843–846. Available at: https://doi.org/10.1242/jcs.01660.