Title: Analysis of context-specific KRAS-effectors (sub)complexes in Caco-2 cells

Authors: Camille Ternet[1,2,†], Philipp Junk[1,2,†], Thomas Sevrin[1,2,†], Simona Catozzi[1,2], Giorgio Oliviero[1], Kieran Wynne[1,3], Christina Kiel[4,1,2*]

Affiliations:

[1]Systems Biology Ireland, School of Medicine, University College Dublin, Dublin 4, Ireland.

[2]UCD Charles Institute of Dermatology, School of Medicine, University College Dublin, Dublin 4, Ireland.

[3]Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland

[4]Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy

[†]Equal author contribution

*Corresponding author. E-mail: christina.kiel@unipv.it

Abstract: Ras is a key switch controlling cell behavior. In the GTP-bound form, Ras interacts with numerous effectors in a mutually exclusive manner, where individual Ras-effectors are likely part of larger cellular (sub)complexes. The molecular details of these (sub)complexes and their alteration in specific contexts is not understood. Focusing on KRAS, we performed affinity purification (AP)-mass spectrometry (MS) experiments of exogenous expressed FLAG-KRAS WT and three oncogenic mutants ('genetic contexts') in the human Caco-2 cell line, each exposed to 11 different culture media ('culture contexts') that mimic conditions

relevant in the colon and colorectal cancer. We identified four effectors present in complex with KRAS in all genetic and growth contexts ('context-general effectors'). Seven effectors are found in KRAS complexes in only some contexts ('context-specific effectors'). Analyzing all interactors in complex with KRAS per condition, we find that the culture contexts had a larger impact on interaction rewiring than genetic contexts. We investigated how changes in the interactome impact functional outcomes and created a shiny app for interactive visualization. We validated some of the functional differences in metabolism and proliferation. Finally, we used networks to evaluate how KRAS effectors are involved in the modulation of functions by random walk analyses of effector-mediated (sub)complexes. Altogether, our work shows the impact of environmental contexts on network rewiring, which provides insights into tissue-specific signaling mechanisms. This may also explain why KRAS oncogenic mutants may be causing cancer only in specific tissues despite KRAS being expressed in most cells and tissues.

Introduction

The interactome of a cell, like a social network, refers to the entirety of interactions of cellular molecules, in particular protein-protein interactions (PPIs) (*1*). These interactions form a network and impact the spatial protein localization and functional organization of a cell. Networks adapt to internal and external cues by converting the signal in responses to stimuli into a plethora of possible output functions that drive cell fates and phenotypes. PPIs, as the core of signaling networks, impact how signals are transduced, and alterations in cellular networks are often linked to diseases, particularly complex diseases such as cancer (*1*, *2*). Mutations in oncogenes can perturb PPI networks (*3*) when protein catalytic and binding functions are affected resulting in alterations in the proteins' binding interfaces (*4*).

The oncoprotein KRAS is an example of a hub signaling protein, as it is part of a highly interconnected and dynamic network capable of interacting with many other proteins (*5*). Oncogenic mutations in KRAS rewire interactions and signaling pathways (*6*). KRAS belongs to the Ras superfamily of GTPases and acts as a molecular switch that cycles between an inactive guanosine diphosphate (GDP)-bound state and an active guanosine triphosphate (GTP)-bound state. The GTP-bound Ras protein mediates binding to several downstream proteins, thereby controlling essential and diverse cellular processes such as survival, polarization, proliferation, differentiation, apoptosis, and migration (*7*, *8*). It is still enigmatic how Ras does all of it. However, what is known is that a class of proteins, called 'effectors', plays a critical role (*5*, *9*, *10*).

Ras effectors are defined as proteins that bind much stronger (i.e. with higher affinity or lower K_d-value) to Ras•GTP than to Ras•GDP. Their interaction with Ras•GTP relies on a domain with a ubiquitin-like topology of three types: the Ras-binding domain (RBD), the Ras association (RA) domain, or the PI3K_rbd, which will herein collectively be referred to as RBDs. All effector RBDs recognize the same switch regions of Ras•GTP, which results in mutually exclusive binding (*8*, *11*). While the presence of an RBD is a necessary condition to qualify as an effector for Ras•GTP, it is not a sufficient criterion. Indeed, for a total of 56 effectors that contain RBD domains, the binding affinities between Ras•GTP-effector complexes are known (either from experiments or computational predictions) to vary, and some are predicted not to bind at all (Fig. S1) (*5*, *8*, *12-14*).

In addition to affinities between the RBDs and Ras•GTP, protein abundances are important for complex formation. In a previous study we used protein abundances together with binding affinities in a mathematical model to predict the amount of each of the 56 effectors in complex with Ras•GTP in 29 human tissues (*15*). Surprisingly, only nine effectors form significant complexes (≥ 5%) with Ras•GTP in at least one of the 29 tissues

(here referred to as group 1 effectors). These results let us wonder about the relevance of the remaining effectors, some of which are well-established effectors, such as PI3-kinase (PI3K) (*16*). As effectors are generally multi-domain proteins, we reasoned that domains that can transfer effectors to the plasma membrane (PM), where Ras•GTP is localized, can increase the number of complexes formed between Ras•GTP and effectors (*8*). Indeed, seminal work by Kholodenko and colleagues has demonstrated that membrane anchoring of both interacting proteins strongly increases the average lifetime of complexes, i.e. the 'piggyback' mechanism (*17*). Indeed, when we applied the piggyback mechanism to the Ras-effector model, we identified 32 effectors that are predicted to form significant complexes with Ras•GTP only with an additional domain recruited to the PM (here referred to as group 2 effectors) (Fig. S1). These effectors were predicted to be recruited to the PM in response to specific conditions (e.g. inputs/stimuli/growth factors) (*5*) (*8*). The remaining 15 effectors are never predicted to be in significant complex with Ras•GTP and are likely no true Ras effectors (here referred to as group 3 effectors).

Colorectal cancer (CRC) is the fourth leading cause of cancer death worldwide. A statistic update reports that there were 1.8 million new cases worldwide in 2018, with a significant burden shift from the old to the younger individuals (*18*). CRC develops through a complex sequence of processes involving an accumulation of epigenetic and genetic alterations, where one of the major driver mutations appears to be KRAS mutations and specific pathways that regulate cell growth and differentiation (*19*). The most frequent KRAS mutations found in CRC are single-point mutations found at codon 12 with the G12D, G12V, and G12C mutations, often followed by codon G13 and Q61 (*20*, *21*). Oncogenic KRAS leads to an accumulation of constitutively active KRAS mutated proteins leading toward the activation of diversified downstream signaling pathways such as the RAS/RAF/MEK/ERK signaling pathway and the PI3K/AKT signaling pathways which were extensively studied in

RAS effectors cancer context (*22*). However, there is evidence that other RAS effectors play a role in cancer (*23*).

In this work, we experimentally probed context-specific network rewiring of KRAS exogenously expressed with a FLAG-tag in immortalized human Caco-2 cells. This cell line, derived from human colorectal adenocarcinoma cells, harbors somatic APC mutations and CTNNB1 (i.e., β-catenin) mutations, but is wildtype (WT) for KRAS (*24*). To probe different 'genetic contexts', we exogenously expressed KRAS WT and three oncogenic mutations frequently found in CRC (G12V, G12D and G12C) to probe different 'genetic contexts'. To probe different 'culture contexts' we grew Caco-2 cells in various growth media mimicking tumor microenvironments (TME) that are known to impact CRC maintenance, progression, and metastasis, which have been described earlier in connection with oncogenic KRAS. Interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α), both part of the inflammatory response found in the TME, are factors of those growth culture contexts (*25-27*). (Patho)physiological conditions such as hypoxia (*28*, *29*) (mimicked by Dimethyloxalylglycine, DMOG, (*30*)), epidermal growth factor (EGF), and prostaglandin E2 (PGE2) also play a role in CRC and KRAS TME (*31-33*) and were selected as growth conditions here. Each combination of genetic and culture contexts was analyzed separately in affinity purification-mass spectrometry (AP-MS) experiments (*34*, *35*) to determine KRAS-mediated complexes. Our study provides an in-depth reconstruction of PPI networks mediated by oncogenic KRAS-effector proteins in specific relevant (patho)physiological colon contexts. Additionally, by identifying different levels of network organization called sub-complexes, we further detailed the downstream pathways mediated by effectors of KRAS and linked them to functional outputs.

Results

Analysis of KRAS-mediated networks in different genetic and culture contexts

We conducted AP-MS experiments to characterize the PPI landscape of both the WT and oncogenic mutant forms of KRAS in different growth conditions. KRAS protein variants were exogenously expressed as FLAG-tagged proteins under the control of a doxycycline-inducible promoter (*36*). As observed earlier (*36*), the promotor shows some leakiness even without doxycycline. As we aimed to express FLAG-KRAS at relatively physiological levels, doxycycline was only added to express the FLAG-KRAS WT proteins at a dose that resulted in equal expression levels compared to the FLAG-KRAS mutant proteins expressed without doxycycline (Fig. S2).

To analyze the KRAS WT and mutant interactomes in different growth media ('culture contexts') that mimic conditions relevant in the colon and CRC, Caco-2 cells were grown 4 h after transfection for 24 h in minimal medium (DMEM with 2 mM L-glutamine) supplemented with either IL-6, TNF-α, PGE2, epidermal growth factor (EGF) or the HIF-hydroxylase inhibitor DMOG at different concentrations (20 and 200 ng/ml) before the AP-MS experiment was conducted. Altogether, we tested four 'genetic contexts' (FLAG-KRAS WT, G12V, G12D and G12C) and 11 'culture contexts' (minimal medium, and each two concentrations of IL6, TNFα, PGE2, EGF, and DMOG in minimal medium), resulting in 44 condition-specific AP-MS experiments (Fig. 1).

To identify high-confidence interacting proteins for each condition, the label-free quantification (LFQ) intensities data for all proteins were filtered in a series of steps (see Materials and methods). Specifically, data for each MS run (44 x three biological with two technical replicates = 264) was visualized as histogram to filter out 8 runs with very few proteins identified (Fig. S3). Further, for a protein to qualify for the high-confidence list, it had to be detected in at least 60 % of the technical and biological replicates for a specific condition. In addition, only proteins that were significantly enriched compared to the beads

only control were included. Technical replicates were merged using the median LFQ intensity. To verify the robustness and applicability of the protocol, we analyzed KRAS expression levels in the complete dataset. The LFQ intensity of the KRAS bait is comparable across all AP-MS conditions (Fig. S4), which suggests that a similar quantity of FLAG-KRAS proteins binds to the magnetic beads across all APs.

A total of 2265 high-confidence PPIs were identified in the 44 contexts, with an average of 725 PPIs per condition (Fig. 1A). Although KRAS is a small protein, a large number of interaction partners is not too surprising, as many of those are expected to be not direct binary physical interactors but rather bind via third proteins (i.e. effectors or other proteins that enable compatible complex formations). Of note, less interactors are generally found for conditions in minimal medium (cf. unstimulated conditions in Fig. 1A), supporting our initial hypothesis that microenvironmental contexts play a significant role in KRAS complex formations. Further, the FLAG-KRAS WT APs generally have less PPIs, which can likely be explained by the fact that effectors bind KRAS predominantly in its GTP-bound form (in fact, no effectors are detected in any of the FLAG-KRAS WT APs grown in minimal medium). Comparing the 2265 high-confidence PPIs determined in this work to 811 previously reported KRAS PPIs (reviewed in (*5*)) shows an overlap of 245 proteins (hence, 30.2 % of the literature PPI are among the 2265 identified here) and 2020 proteins were not previously reported (Fig. 1B). Similar overlaps are obtained when focusing only on the classical effector proteins (Fig. 1C).

To gain insights into the whole dataset, a principal component (PC) analysis and a uniform manifold approximation and projection (UMAP) (*37*) were performed with all high-confidence interactors identified in each AP-MS experiment (Fig. 1D; Fig. S5). Both techniques enable a dimensionality reduction of the data and a data visualization. The PC analysis, which is commonly used, tries to preserve the global structure of the data (Fig. 1D),

while the UMAP tries to preserve the data's local structure (Fig. S5). The unstimulated and KRAS WT samples cluster separately from the other groups suggesting that KRAS WT and unstimulated conditions are a good control group (KRAS mutants and stimulated conditions; middle left area in Fig. 1D). Interestingly, KRAS interactor proteins detected in the different mutant datasets cluster together, compared to the different culture context datasets, where the data are more discriminated. For example, IL-6 and PGE2 context cluster together at the top right corner, whereas DMOG context at the bottom right corner (Fig. 1D). Taken together, these results suggest that the proteins detected in complex with KRAS seem to be more conditions-dependent rather than mutation-dependent.

Binding landscape of effectors in complex with KRAS WT and mutants

Effector proteins bind Ras in the GTP-bound state and they are likely forming, among other proteins, the first layer of interacting proteins. Hence, we first characterized the KRAS-effector layer in more detail. As previously mentioned, no effectors are found in complex with KRAS WT in minimal medium, which is expected as KRAS will be mainly in the GDP-bound state that does not enable high affinity binding. Eleven out of 56 classical Ras effector proteins were identified in at least one of the 44 conditions (Fig. 2AB). All effectors identified in complex with KRAS belong to either group 1 (AFDN, ARAF, RAF1, BRAF, RGL2) or group 2 effectors (RIN1, PIK3CA, GRB7, RIN2, PIK3C2A, ARAP1). They are generally highly expressed in colon tissue and Caco-2 cells (with medium or high affinities for Ras•GTP) or are moderately expressed but have high affinities in complex with Ras•GTP (Fig. 2B). Concerning the 45 effectors not found in any of the KRAS AP-MS samples, 15 belong to group 3 effectors (likely no “true” Ras effectors) and 26 belong to group 2 (of which 7 have low mRNA levels in Caco-2; <3.3 nTPM). Four effectors belong to group 1 effectors, of which RALGDS and RASSF5 are part of the KRAS literature interactome

literature and are highly/moderately expressed in Caco-2 cells. SNX27 and RASSF7 are also highly/moderately expressed in Caco-2 cells, but their affinities in complex with Ras are lower. Altogether, we provide a near-to-complete binding landscape of effectors in complex with KRAS under the conditions tested.

A comparison with the effectors identified in previous KRAS interactomes (reviewed in (5)) (Fig. 2CD) shows that the AP-MS experiments in this study specifically increase the percentage of group 2 effectors, but little increase in group 1 effector coverage and no increase in group 3 effectors (Fig. 2E). Further, the number of conditions in which an effector is identified in complex with KRAS tends to be lower for group 2 effectors (Fig. 2F). To visualize in which genetic and culture contexts the 11 effectors were detected, two heatmap images were generated (Fig. 3). The two heatmaps show a similar pattern in terms of effector detection and abundances in each of the groups of the AP KRAS mutants (i.e., G12D, G12C and G12V) in unstimulated and stimulated conditions. More specifically, the effectors AFDN, ARAF, RAF1 and RIN1 are detected in all the KRAS mutant AP-MS experiments when unstimulated and stimulated (Fig. 3A). These effectors are also detected in the WT AP-MS experiments – albeit not in all the culture contexts. They are classified in the effector group 1 except for RIN1, which is part of the effector group 2. Moreover, they appear to be more abundant when detected in particular conditions such as DMOG and TNF-α, compared to other conditions such as IL-6 or PGE2 (Fig. 3B). As these effectors are detected consistently in the presence of KRAS with or without stimulations, we propose that these effectors are KRAS-specific rather than condition(stimulation)-specific. Other effectors are only detected in specific stimulated conditions and are mainly detected in the predicted effector group 2. Moreover, GRB7 is only detected in the presence of DMOG culture context. Another effector, PIK3CA, is only detected when stimulated with TNF-α in the presence of either the G12D or G12V KRAS mutations. The effector PIK3C2A is only detected

significantly in the presence of the G12D KRAS mutation with DMOG. These effectors grouped in effector group 2 can be classified as conditions-specific. To mention, the effector BRAF, which was computationally predicted to be always in complex with KRAS, is found in complex to KRAS only in DMOG and TNF-α conditions.

Altogether, this supports our initial hypothesis that group 2 effectors tend to be found in complex with Ras only in specific conditions that promote PM recruitment via RBD-independent mechanisms (Fig. S6). Further, it validates our computational-based classification into group 1, group 2, and group 3 effectors (*15*) and supports its applicability beyond the 29 human tissues as the basis of the prediction model.

Investigation of functional differences in the KRAS interactome

To investigate functional differences in the interactome of the different genetic and culture contexts, two approaches were chosen. First, a differential interaction analysis was performed on the identified proteins followed by a gene set enrichment analysis (GSEA) against the Gene Ontology (GO) Biological Processes gene ontology (Fig. 4A). Secondly, using the ontology, we collapsed each sample onto all ontologies listed under “biological process” by summing up their LFQ intensities. This process preserved the variation in the data (Fig. S7). Then, multiple ANOVAs were used to identify differences between the samples in terms of their summed up LFQ intensities. Out of the 16.000 GO terms tested, we find significant changes for 2135 (Fig. 4B; Fig. S8A). Afterwards, semantic analysis was used to organize the significantly changed ontology terms into clusters. The full distance heatmap from the semantic analysis together with the clusters and some of the data projected onto it is shown in Fig. 4B.The biggest semantic clusters are linked to metabolic and biosynthetic processes (cluster 1), signaling and immune response (cluster 2), vesicle-mediated transmembrane and

ion transport (cluster 3), differentiation, development and morphogenesis (cluster 4), and actin and cytoskeleton organization (cluster 5) (Fig. 4B, Fig. S8D-H). Smaller clusters are linked to thermogenesis, ion homeostasis, cell cycle, leukocyte activation, regulation of GTPases, proliferation, and apoptotic cell death (Fig. 4B). Altogether, the overall functional differences in the KRAS interactome are consistent with known cellular functions mediated by KRAS (*8*, *38*).

The above two-fold analysis shows different aspects of functional differences between the different genetic and culture conditions. To make the data better approachable, we developed an interactive R Shiny app for exploring the functional terms that are different between the samples (Fig. S9). Users can explore the analysis through the semantic distance heatmap and the semantic clusters, or search and filter for functional terms of their interest, visualize which proteins are part of this particular GO term and show their abundance in the different samples. It also directly displays the samples that are statistically significant to each other. We propose this app as a resource to filter for interesting functional influence of certain KRAS mutations or certain growth conditions from our data set. From the results of this functional investigation, we selected some GO terms of interest to us which we went on to validate in the wet lab. Those were GO terms related to proliferation ("Epithelial cell proliferation" and "Positive regulation of cell population proliferation"), glucose metabolism ("Glycolytic process" and "Regulation of glucose metabolic process"), and ATP metabolism ("ATP metabolic process", and "Regulation of ATP metabolic process") shown for the genetic contexts G12C and G12D and the culture contexts DMOG and IL-6 in Fig. 5A). As we are studying the effect of oncogenic KRAS-mutation in an adenocarcinoma cell line under CRC microenvironment mimicking culture contexts, we were particularly interested in biological processes commonly observed in cancer development. Among the hallmarks of cancer (*39*), sustained cell proliferation and the Warburg effect, described by a switch in cell

energetic metabolism from oxidative phosphorylation to aerobic glycolysis, are two biological and metabolic processes that are feasible to test experimentally. In the GO functional analysis, we observed that PGE2 and DMOG had the greatest number of differentially expressed GO compared to unstimulated while IL-6 had the lowest. In addition, we observed that the differential expression profile of the three KRAS mutants were similar, and KRAS G12C and G12D differ the most for GO terms related to biological and metabolic processes.

Analysis of cell phenotypes in selected genetic and culture contexts

To investigate changes in cell proliferation and glycolytic metabolism for three genetic contexts (KRAS WT, G12D and G12C) in three culture contexts (unstimulated, DMOG and IL-6), we measured cell count, cell viability, glucose uptake and lactate release over a 72 h period in Caco-2 cells (see methods). We chose to determine Caco2 metabolism and proliferation until 72 h post-transfection under the assumption that the changes observed in the interactome seen 24 h post-transfection would affect cell phenotype on a long-term basis and that phenotypic changes can take longer to be accurately measured. To ensure that exogeneous KRAS was expressed until 72 h post-transfection we performed a Western-Blot of FLAG-KRAS before transfection and from 24 h to 72 h post-transfection. We observed an overall significant increase of FLAG expression 24 h post-transfection with a peak at 48 h, and then a decrease of FLAG expression at 72 h to reach the level of 24 h time point (Fig. S10). Those results suggest that for all genetic contexts, exogenous KRAS was effectively expressed during a 72 h period. However, we also noticed that exogenous KRAS expression was significantly different between the three genetic contexts.

We observed a significant increase in cell proliferation under DMOG context compared to the control (unstimulated), although IL-6 led to a significant decrease in cell proliferation (Fig. S11A). The increase in cell proliferation in DMOG culture context was confirmed by cell viability results (ATP concentration, Fig. S11G). Furthermore, for those two phenotypic parameters, under DMOG stimulation, KRAS G12D had a significantly greater cell proliferation than KRAS WT and G12C. In terms of glycolytic metabolism, it was overall better captured by lactate release than glucose uptake and the highest lactate release was observed in DMOG context. A greater lactate release was also observed for both KRAS mutants compared to WT (Fig. S11E). When normalized per cell, both glucose uptake and lactate release were significantly greater in the IL-6 condition than in the control and DMOG, suggesting a higher glucose utilization at the single cell level in IL-6 context. Finally, the ATP pool per cell, which can be used as a proxy for changes in metabolic activity, was higher for both stimulations compared to the control and for both oncogenic KRAS compared to the WT. Altogether, the experimental results suggest that under genetic and culture contexts that mimic colon and CRC, Caco-2 cells produce more energy than KRAS WT cells in control conditions. It also suggests that changes on the level of the interactome, particularly those seen with different culture (microenvironmental) contexts, likely manifest as phenotypical changes. It supports our initial hypothesis that culture contexts are modifiers of both KRAS-mediated networks and downstream phenotypes.

Next, we compared the measured phenotypical parameters to the sum of LFQ intensities of proteins associated with GO terms linked to proliferation, glucose metabolism, or ATP metabolism with either term directly related to the final cell phenotype or its regulation (Fig. 5). Overall, the culture context effects are similar between the AP-MS (GO term) and the phenotypical experiments. Indeed, we observed that for all GO terms excepts “Glycolytic process” the sum of LFQ intensities were significantly higher in DMOG

compared to unstimulated and IL-6 (Table S1). Those results are similar to the cell density, lactate release, and ATP concentration experimental results. In addition, for "Regulation of glucose metabolic process" the sum of LFQ intensity was higher in the IL-6 context compared to unstimulated, which could be linked to the higher lactate release per cell observed in the IL-6 condition. The only notable difference is observed for the "Glycolytic process" where the sum of LFQ intensities was significantly lower in the IL-6 culture context, which is contrary to what was observed for lactate release per cell. Results of the effect of the genetic context are also to some extent similar, with no significant effect on proliferation related GO terms such as cell density. In addition, a significantly higher sum of LFQ intensities was observed in the "Regulation of glucose metabolic process" in the two KRAS mutant proteins compared to WT, with results similar to the lactate release. However, no genetic context effect was observed for ATP metabolism-related GO terms, although mutant KRAS had a higher overall ATP concentration than WT. Especially, the sum of LFQ intensities of "Glycolytic process" was significantly lower in G12C than in the two other genetic contexts mostly due to a substantial decrease under IL-6, which is opposite to what was found for overall lactate release. Those results suggest first that the strongest changes in phenotype, such as those observed for DMOG stimulation, are more reliably captured by the AP-MS than the milder changes such as those observed for genetic context or IL-6 stimulation. Second, the changes observed for GO terms related to the regulation of a process seem to be better at capturing the results of phenotypical experiment than GO terms related to the biological process itself. This latter point is also suggested by the better Pearson's correlation coefficient observed for phenotypical parameters with the associated GO terms related to the regulation of a biological process than those directly related to the biological process (Fig. 5C). This may be because proteins that have a direct effect on phenotype are too many layers downstream of RAS signaling network to be precisely captured by AP-MS

experiments. Finally, for metabolic parameters, the functional analysis results are more correlated with overall lactate release and ATP concentration than with the values normalized per cell. This implies that the results of the functional analysis based on AP-MS data are more suitable at capturing changes in metabolism at the level of the whole cell culture than at the level of a single cell.

Information flow analysis predicts the contribution of effectors to functional processes

After demonstrating that the differences in the summed LFQ intensities for a functional process can be associated with functional differences in the behavior of the Caco-2 cells, we aimed to further explore how the changes in the interactome led to these functional differences. In particular, we were interested in understanding which effectors and proteins downstream of KRAS were involved for a specific process. To this end, we used random walks over a filtered version of the STRING network, in which we biased the random decision for each step depending on whether a potential next protein was part of the AP-MS dataset for a specific sample or not (see methods). This left us with a collection of paths and their probability to be traversed, based on the network architecture and the proteins found in the AP-MS sample.

Applying this method to the functional terms and samples we were interested in, we observed drastic changes in the network architecture depending on the sample (Fig. 6). Particularly, for some GO terms there is a different predicted engagement of effectors. Examples are the two GO terms related to “Epithelial Proliferation” for KRAS oncogenic mutations in culture contexts IL-6, DMOG and unstimulated (Fig. S12). For “Epithelial cell proliferation”, in IL-6 context high path counts are found for the effectors AFDN, ARAF, and RAF1 (Fig. 6A). In DMOG context, additional high counts are found for BRAF, GRB7, and

PIK3CA (Fig. 6A). Likewise, for "Positive regulation of cell population proliferation" contributions of AFDN and PIK3CA dominate in DMOG context (Fig. 6B). For GO terms related to glucose metabolism, AFDN, GRB7 and PIK3CA dominate the path counts for "Glycolytic process (GO:0006006)" in DMOG condition and for KRAS G12C in IL-6 context, but has a low count for "Regulation of glucose metabolic process (GO:0010906)", where ARAF, BRAF and RAF1 dominate in most culture conditions (Fig. 6CD; Fig. S13). With respect to GO terms related to ATP metabolism, the culture contexts IL-6 and DMOG show profound path count differences for the effector RAF1, which contributes more in DMOG than in IL-6 context (for KRAS G12D genetic context), but slightly higher path count in IL-6 context for KRAS G12C (Fig. 6EF; Fig. S14). The path count is generally low for BRAF and ARAF. AFDN has high/ the highest path counts in almost all genetic and culture contexts. Also noteworthy, RIN1 is present in all contexts but not necessarily much involved. Altogether, our biased random walk analyses predicts the contribution of individual effectors to GO terms that link to experimental phenotypes.

Discussion

This study set out to explore KRAS as a key cellular signaling hub in specific relevant (patho)physiological contexts. The Caco-2 cell line has been used as a relevant model system that can be grown in various growth media (culture contexts) and enable exogenous expression of KRAS WT and oncogenic mutants (genetic contexts). Indeed, Caco-2 cells are human intestinal epithelial cells that closely mimic the colon intestinal epithelium in the early stage of CRC. By identifying different levels of network organization (e.g. sub-complexes and number of paths traversing a network), we aimed to detail the downstream pathway of KRAS further and investigate the functional outputs. To address this challenge

experimentally, even though all methodology has their limitations, AP-MS excels for profiling interactomes in humans (*34*) due to its sensitivity and its ability to detect interactions within complexes in appropriate contexts (*40*). We successfully pulled down complexes by using the exogenous expression of tagged bait proteins for different KRAS mutations on Caco-2 cells.

Effectors bind to KRAS in a mutually exclusive fashion and can potentially compete for binding (*9*). Our earlier computational predictions suggested that there is a considerable impact of the culture contexts on the recruitment of specific effectors to the PM (*15*). Here, we identified a total of 11 effectors in at least one of the AP-MS experiments, of which seven are only found in complex with KRAS in some genetic and culture contexts. For example, effectors such as PI3KCA, RIN2, GRB7 or ARAP1, are detected in the presence of culture conditions such as hypoxia (HIF-stabilization), IL-6, and TNF-α. We predict that in these cases the affinity between the RBD and KRAS is not high enough to allow for sufficient binding and that additional domains present in effectors are required to increase the number of complexes formed between KRAS and effectors at the PM (based on the "piggyback" mechanisms; (*17*)). Indeed, we show in this work that the total number of effectors and other proteins in the Ras-mediated complex increases with the number of conditions. This context dependent binding can be explained by the fact that cells in their physiological microenvironment are constantly experiencing a variety of stimuli that trigger receptors (e.g. the EGF receptor) located on the plasma membrane, where Ras is located (*41*). In the context of cancer, tumor cells are often located in a hypoxic, immunosuppressive and nutrition-deficient microenvironment that causes reprogramming of metabolism and signaling (*39*). Indeed, we identified culture context-specific metabolic alterations in glucose and ATP metabolism in the Caco-2 cells. Hence, this work supports the requirement to study the role of the microenvironment when performing an experiment that aims to characterize PPI

networks because they have a major role in driving complex formation. It also demonstrates the need to consider multidomain interactions. However, the interpretation of PM recruitment and culture conditions might not be straightforward. In fact, it is difficult to understand and predict what happens on the upstream level of Ras and effectors and why some effectors are identified in the specific conditions tested. This is partly due to signaling pathways that are highly cell type-specific (*2*, *42*). Together with different databases such as the Human Protein Atlas initiative (*43*) and the large-scale interactome Bioplex based on AP-MS baits (*44*, *45*), analyzing human PPI and the conditions in which they occur is essential for the development of a context-dependent human interactome. Indeed, a new version of the Bioplex 3.0 interactome has been recently published where, in addition to the HEK 293T cells, a dual comparison with the HCT116 cell line was performed (*40*).

Based on the assumption that effectors compete for binding to KRAS, our working hypothesis is that individual KRAS-effector-mediated sub-complexes form in a cell, which ultimately affect downstream signal propagation and cellular phenotypes. Indeed, we show here that differences in KRAS-mediated complexes propagate to downstream changes in phenotypes that roughly align with the predicted functional changes based on GO terms of proteins detected in an AP-MS experiment. This suggests that the PPI network orientation/ assembly on the level of KRAS (likely mediated in part by effectors in complex with KRAS) impacts the downstream phenotype.

As effectors compete for binding to KRAS we hypothesized that specific Ras-effector sub-complexes exist that each (or in combination) link to specific phenotypes. To explore the contribution of individual effectors to phenotypes, we used biased random walk analysis. Indeed, we find differences in the number of paths between different genetic and culture contexts. The analysis also enabled us to predict which effector pathways are likely linked to cellular phenotypes. Hence, our analysis pipeline that combines AP-MS data with random

walks and GO terms provides a novel way to link PPI networks to phenotypes. The pipeline and code is available to the scientific community and can be adapted for specific AP-MS experiments. There are, however, limitations of the random walk analysis. The data structure at the end is a collection of different paths for different targets for different conditions. Some of these are comparable, some are biased. Paths ending into the same target should be comparable across conditions, as long as the underlying network structure does not change. However, there is a bias for shorter paths to be more likely to be found, and targets with shorter shortest-paths have on average higher counts in the found paths. Additionally, the analysis is strongly dependent on the underlying network structure that is used.

With respect to the impact of genetic vs culture contexts on KRAS-mediated network rewiring, our analyses based on PCA and UMAP suggest that the impact of growth condition (culture context) is greater than type of oncogenic mutation (genetic context). We observe a similar trend in the functional analysis as well as in the effector contributions as calculated by random walks, where different oncogenic mutants generally having more similar effector path counts are different culture contexts/ growth conditions. Indeed, the results of this work offer an additional explanation why cancer genes and mutations only manifest in some but not all tissues (*46*).

Future steps in systems medicine require the integration of protein abundances with context-specific conditions and localized signaling responses. Indeed, quantitatively predicting the influence of specific conditions on larger networks to get an efficient predictive model would be ideal, especially in the case of oncogenic mutations. In addition, understanding the rewiring in physiological contexts to enhance the understanding of network rewiring in cancer contexts would provide new insights into potential therapeutic targets (*47*).

Materials and Methods

Culturing of Caco-2 cells

Caco-2 cells (ATCC©HTB-37) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco™, ThermoFisher, 21969-035) supplemented with 2 mM L-glutamine (Gibco™, ThermoFisher, 25030-024), 10 % (v/v) Foetal Bovine Serum (FBS) (Gibco™, ThermoFisher, A4766801), and 1 % Penicillin/streptomycin (Gibco™, ThermoFisher, 15140122). For long-term storage, frozen stock vials were made on the week of receiving the cell line in Recovery™ cell culture freezing medium (Gibco™, ThermoFisher, 12648010) and stored in liquid nitrogen. For each experiment, cells were not exceeding passage 25 and were thawed from the liquid nitrogen stock. To generate growth media that mimic conditions relevant in the colon and CRC ('culture contexts'), the minimal medium (DMEM with 2 mM L-glutamine) was supplemented with either IL6 (interleukin-6) (ThermoFisher), TNFα (tumor necrosis factor-alpha) (ThermoFisher), PGE2 (prostaglandin E2) (ThermoFisher), EGF (epidermal growth factor) or the HIF-Hydroxylase Inhibitor DMOG (Cayman chemical) at different concentration (20 and 200 ng/ml).

Plasmids for exogenous expression of FLAG-KRAS WT and mutants

Plasmids were gifted from the previous research laboratory of Christina Kiel in Barcelona (CRG) (from Luis Serrano and Hannah Benisty). All the plasmids harbor the identical backbone pMDS-TetOn3G-kozak-FLAG-GOI (gene of interest). Plasmids differ only by their GOI, which are wildtype (WT) KRAS, KRASG12D, KRASG12V, or KRASG12C as GOI.

Bacterial transformation with plasmids, plasmid extraction and purification

The bacterial transformation of the plasmids for exogenous expression of FLAG-KRAS (pMDS-TetOn3G-kozak-FLAG-KRAS WT/mutants) was performed using the One-Shot™ Stbl3TM (Invitrogen, C737303) chemically competent bacterial cells to replicate each plasmid following the manufacturer's instructions. Subsequently, 100 µl of the bacteria-plasmids solutions were plated into LB selective agar plates containing the antibiotic spectinomycin (50 µg/ml). Plates were incubated at 37 °C, overnight. The next day, individual bacterial colonies were selected from LB agar plate and grown in 4 ml LB broth with the corresponding antibiotics for 6 to 12 h at 37 °C in a shaking incubator at 250 rpm. After incubation, several aliquots of this original starter culture were used to generate a bacterial glycerol stock for long-term storage at -80°C (1 ml transformed bacteria in 1 ml 50% glycerol). The remainder of the original starter culture was then used to grow at a large scale the transformed bacteria under selective antibiotics overnight in 500 ml of LB medium at 37°C in a shaking incubator at 250 rpm. The HiSpeed Plasmid Maxi Kit (Qiagen) was used to generate larger amount of the FLAG-KRAS plasmids. The kit was used following the manufacturer's instructions. Final DNA was eluted in 400 µl of TE buffer and allowed to resuspend overnight to ensure homogeneity. The following day, concentrations and purities were measured on the Implen NanoPhotometer® NP80, and plasmids DNA were stored at -20°C.

Transfection and expression of FLAG-KRAS in Caco-2 cells

For AP-MS experiments, Caco-2 cells were seeded 24 hours before transfection in 10-cm dishes in normal growth medium and grown to 70-80% of confluency. Cells were transfected with 15 µg of pMDS-TetOn3G-kozak-FLAG-GOI plasmids (containing FLAG-KRASWT or

FLAG-KRASG12D or FLAG-KRASG12V or FLAG-KRASG12C as GOI) using Lipofectamine 2000 (Invitrogen, 11668-019) according to the manufacturer's instructions in OPTI-MEM reduced serum medium (Gibco™, ThermoFisher, 31985-062) for 4 h. Then, the medium was changed and supplemented with culture medium containing the various growth conditions. To note, cells transfected with the KRASWT plasmids were always supplemented with 15ng/ml of doxycycline (Sigma-Aldrich). Cells were incubated for 24 h at 37°C and harvested. FLAG-KRAS mutant plasmid transfections were not supplemented with doxycycline as the promoter is leaky and KRAS mutant proteins were already expressed at WT levels without adding doxycycline.

Caco-2 cell lysis, protein extractions and concentration

Caco-2 cell lysates were obtained after trypsinization, and cell pellets were recovered and washed twice with PBS 1X. The cells pellets were then resuspended in the appropriate volume (e.g., 300µl for the AP-MS experiments) of lysis buffer (50 mM TRIS HCL pH 7.5, 1 mM EDTA, 1 mM EGTA, 150 mM NaCl, 2 mM MgCl2, 1 mM DTT, and 1 % IGEPAL/NP-40 supplemented with PhosSTOP (Roche) and cOmplete, Mini protease inhibitor cocktail (Roche)). Cells were lysed for 30 minutes on a rotator at 4°C, centrifuged at 14000 rpm for 30 minutes at 4 °C, and the supernatants were collected in a new tube. Protein concentrations were measured using the Pierce™ 660-nm Protein Assay (ThermoFisher, 22660) per manufacturer's guidelines. Samples were incubated for 5 min before absorbance was read at 660 nm on a SpectraMax M3 plate reader. Net absorbances were plotted against BSA protein concentration for standard curve generation (ThermoFisher, 23208). For each sample, the concentration was obtained by comparing net absorbance values against the generated standard curve. A new standard curve was generated for each assay. Kept on ice, cell lysates were then directly used for affinity purification.

Western blotting

Prior to loading the samples into the gel, a normalization of the concentration for each sample is done, with a concentration aiming to be 1ug/ul. Proteins were then denatured by incubating samples at 95°C for 5 min in 4 x Laemmli buffer and dithiothreitol (DTT) before loading onto 4 to 12% NuPAGE gradient precast gels (ThermoFisher). Gels were run for 10 min at 110 V, followed by 45min at 150 V, with gels submerged in NuPAGE MES running buffer (ThermoFisher). After electrophoresis, proteins are dry-transferred using the iBLOT2 device (ThermoFisher) for 7 min into a nitrocellulose membrane. The membranes were checked by Ponceau S staining to ensure protein transfer. Then, the membranes were washed in 1 X Tris buffer saline-tween20 (TBS-T) before blocking solution for 1 h in 5 % milk at room temperature in a shaking device. Depending on the antibody, the membranes were incubated either overnight at 4 °C or at room temperature for 4 h, with the primary antibody diluted in 0.05 % milk in TBS-T. The membranes were then washed three times in TBS-T, 10 min each and incubated horse radish peroxidase (HRP) conjugated secondary antibody for 1 h diluted in milk TBS-T on a shaking device. Protein bands were developed using high sensitivity ECL reagent (ThermoFisher) with the West Pico western blotting substrate per the manufacturer's instructions and visualized using the G-Box image developer (SYNGENE). Densitometry analysis was performed using ImageJ, with target protein bands normalized to loading control (beta-actin or GAPDH). The following antibodies were used for Western blotting: beta-Actin (Cell Signaling, #4970, rabbit / monoclonal, 1/3000 dilution), GAPDH (Abcam, ab2118, rabbit / monoclonal, 1/1000 dilution), panRAS (Abcam, ab52939, rabbit / monoclonal, 1/5000 dilution), KRAS (CPTC-KRAS4B-2, DSHB, mouse / monoclonal, 0.5 ug/ml working concentration), and secondary Anti-mouse HRP (Abcam, ab97023, goat / monoclonal, 1/3000).

Affinity purification (AP)

Caco-2 cell lysates expressing FLAG-KRAS proteins were immunoprecipitated from 800 µg of cell lysate using anti-FLAG-M2 magnetic beads (Sigma, M8823) by using the KingFisher DuoPrime purification system (ThermoFisher). Beads were washed in TBS (according to the manufacturer's instructions) for 5 min, twice, at low speed. Then beads were collected by the KingFisher magnet and discarded into the samples wells and mixed at a slow speed for 1 hour. Beads-antibody samples were collected and went through different wash salted solutions (Wash 1 and 2: RIPA buffer with 150 mM NaCl; Wash 3: RIPA buffer with 500 mM NaCl), mixed at low speed for 30 seconds. Beads-antibody-sample were eluted in 50 µl of glycine (0.1 M, pH 3.0) for 5min. Immediately after, samples were neutralized with 20 µl of TRIS BASE (1 M, pH 8.0).

Sample preparation after AP for mass spectrometry (MS)

For protein cleanup the paramagnetic bead–based SP3 (solid-phase-enhanced sample-preparation) workflow was used (*48*). For each AP experiment sample protein concentrations were determined using the Pierce^{TM} BCA protein assay (ThermoFisher) following the manufacturer's instructions and 50 µg of proteins were adjusted in 20 µl of buffer/MS grade water. Samples were homogenized and denatured in urea (final concentration, 4 M), ammonium bicarbonate (100 mM), and calcium chloride (100 mM), then reduced in DTT (final concentration, 1 mM) for 15 minutes at room temperature and alkalinized in iodoacetamide (IAA) (3 mM) in the dark at room temperature for 15 minutes. The tryptic digestion protocol was performed using the KingFisher DuoPrime purification system (ThermoFisher) in a series of steps. First, magnetic hydrophobic and hydrophilic beads were

washed several times in MS grade water and added to the deepwells plate in the KingFisher along with the samples and the same volume as the sample of 100 % ethanol. Next, the solutions were mixed at low speed for 10 minutes, after which the beads coupled to the proteins were collected with the magnetic arm of the KingFisher and transferred to be washed in 3 different deepwells containing each 80% of ethanol. The washed beads-proteins were then released into the trypsin (Promega, V5111)-containing deepwells at a 50:1 (w/w) protein to protease ratio and mixed at low speed for 8 h of digestions into peptide fragments at 37 °C in the KingFisher. Peptide samples were transferred into low protein binding tubes, 1% of trifluoroacetic acid (TFA) was added to acidify the samples ready to be desalted, cleaned, and concentrated on C18Tips (ThermoFisher, 87784) (*49*) according to the manufacturer's instructions. Purified peptides were dried and resuspended in low protein binding tubes before mass spectrometry analysis in 30 ul of 0.15 % TFA acid and 1 % acetic acid in mass spectrometry grade water.

Mass spectrometry

The peptides were analyzed using a MS shotgun proteomics technique. This technique allows a sensitive bottom-up approach that consists of separating peptides resulting from protein digestion by liquid high-performance liquid chromatography (HPLC) followed by tandem mass spectrometry (MS/MS). Samples were run on a Bruker timsTof Pro mass spectrometer connected to an Evosep One liquid chromatography system. Tryptic peptides were resuspended in 0.1 % formic acid and each sample was loaded onto an Evosep tip. The Evosep tips were placed in position on the Evosep One, in a 96-tip box. The autosampler is configured to pick up each tip, elute and separate the peptides using a set chromatography method (*50*). The chromatography buffers used were buffer B (99.9 % acetonitrile, 0.1% formic acid) and buffer A (99.9 % water, 0.1 % formic acid). All solvents are LCMS grade.

The mass spectrometer was operated in positive ion mode with a capillary voltage of 1500 V, dry gas flow of 3 l/min and a dry temperature of 180 °C. All data was acquired with the instrument operating in trapped ion mobility spectrometry (TIMS) mode. Trapped ions were selected for ms/ms using parallel accumulation serial fragmentation (PASEF). A scan range of (100-1700 m/z) was performed at a rate of 5 PASEF MS/MS frames to 1 MS scan with a cycle time of 1.03 seconds (*51*).

The data analysis was done using MaxQuant software (*52*). The raw data was searched against the Homo sapiens subset of the Uniprot Swissprot database (reviewed) with the search engine Maxquant (release 2.0.3.0). Specific parameters for trapped ion mobility spectrometry data-dependent acquisition (TIMS DDA) were used: Fixed Mod: carbamidomethylation; Variable Mods: methionine, oxidation; Trypsin/P digest enzyme (maximum 2 missed cleavages); Precursor mass tolerances 10 ppm; Peptide FDR 1 %; Protein FDR 1 %. The normalized protein intensity of each identified protein was used for label-free quantitation (LFQ) using the MaxLFQ algorithm (*53*).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*54*) partner repository with the dataset identifier PXD035399.

AP-MS data filtering and ID mapping

The data were first filtered based on the label-free quantification intensities (LFQi) using the following five steps: (i) removal of proteins that were labeled as "only identified by site", "potential contaminant", and "reverse"; (ii) removal of all observations with LFQi equals to 0; (iii) removal of outlier samples (based on low overall LFQi; see Fig. S3); (iv) removal of proteins which are not present in at least 60% of the samples of a group for each group (a group is defined as the collection of three biological with 2 technical replicates for one

condition, which results in a group size of maximum 6); (v) filtering against the negative control sample, which is only the beads used for the AP-MS sample preparations, by only considering proteins for further analysis that are significantly higher found in the samples compared to the negative control. In MS analysis based proteomic data, there are typically two types of missing values, the missing not at random (MNAR) and the missing at random (MAR) (*55*). A mixed imputation strategy was chosen, with kNN imputation as the strategy for MAR values (*56*, *57*) (*58*). Other missing values were considered MNAR values and imputed at value 0. After the imputation, differential interaction analysis was performed for each group against the beads control. P values were adjusted using FDR correction as described by Benjamini-Hochberg (*59*). Afterward, all proteins were extracted for each group which were significantly enriched in the sample (cutoffs: p-value adjusted: < 0.01, Log Fold Change: > 1). The data were transformed to have consistent protein and gene names annotations following the data filtering. The data are received from the MaxQuant software in UniProt IDs and mapped to HGNC gene names using the HGNC database (retrieved 12/2021) However, one UniProt ID can correspond to multiple HGNC gene names. In this case, manual selection of the gene names of interest was performed. Finally, the HGNC names were mapped to genes IDs of the SysGO database (*60*). A couple of proteins could not be found in the SysGO database, and one protein was renamed (i.e., HGNC name: PHB1, which was renamed PHD for SysGO). Then, the technical replicates were merged using the median. In summary, we obtain a dataset with raw LFQi (Data S1) or log2 transform (Data S2) data with biological triplicates. Data preparation was performed in R (http://www.r-project.org/index.html) using the following packages: dplyr (*61*), tidyr (*62*), stringr (*63*), tidyxl, purr (*64*), DEP (*65*) and limma (*66*, *67*). The script file for the data preparation as well as the data pre and post preparation are available on Zenodo (https://zenodo.org/record/6896565#.Yt6m63ZByUk) (DOI: 10.5281/zenodo.6896565).

Functional analysis of the interactome

Functional analysis of the interactome was performed in two different ways. The first approach consists of a differential interaction analysis based on the filtered LFQ intensities. Imputation was performed by a mixed imputation strategy, using bpca (Bayesian PCA) (*68*, *69*) for MNAR values and MinProb (https://cran.r-project.org/web/packages/imputeLCMD/index.html) for MAR values. Differential analysis was performed using limma (*66*, *67*) and DEP (*65*). P values were adjusted using FDR correction by Benjamini-Hochberg (*59*). The results of the differential interaction analysis were evaluated for functional enrichment by performing a gene set enrichment analysis (GSEA) using ClusterProfiler4 (*70*, *71*) against the Gene Ontology (GO) Biological Process (BP) ontology (*72*, *73*).

For the second approach, LFQ intensities were collapsed on GO BP terms by summing up all intensities of all identified proteins for each sample for each GO term. Then, for each GO term, a three-way ANOVA was performed with the main effects of mutation status (genetic context), condition and concentration (culture context) and their interaction terms. The p-values of these ANOVAs were collectively corrected using correction by Holm (*74*). After correction, significant terms ($p < 0.05$) were further analyzed using Tukey's Honest Significant Differences post hoc tests. P-values were collectively corrected using FDR correction by Benjamini-Hochberg (*59*).

Both approaches identify many GO terms that are significantly different ($p < 0.05$ after respective adjustment) between the groups. In order to gain an overview over the results, semantic similarity between the GO terms was determined using the methodology proposed by Schlicker et al. (*75*, *76*). Based on the resulting similarity matrix, GO terms were clustered

using the binary cut algorithm (*77*). The results were visualized as a heatmap with data from the analysis projected as additional heatmaps (*77*, *78*).

All analysis in this part was performed using the R programming language and the tidyverse environment (*62*). The scripts and output for this analysis are available on Zenodo (https://zenodo.org/record/6896565#.Yt6m63ZByUk) (DOI: 10.5281/zenodo.6896565).

Visualization of AP-MS data in shiny app

The results from the functional analysis together with the filtered AP-MS data were put together in an R shiny dashboard, allowing the interactive exploration of our analysis and data (*79*, *80*). The R shiny app is available at https://pjunk.shinyapps.io/kras_apms_vis/, with the source code and underlying data files available at https://github.com/PhilippJunk/kras_apms_vis.

Assessment of phenotypic and metabolic parameters of Caco-2 cells

Caco-2 cells cultured in DMEM supplemented with 10 % FBS were seeded at about 70 % confluency in nine 12-well plates (CELLSTAR, Greiner Bio-One) to test 3 KRAS mutant status (WT, G12D and G12C) in three contexts (unstimulated, DMOG and IL-6). Twenty-four hours post-seeding cells were transfected with FLAG-KRASWT, FLAG-KRASG12D or flag-KRASG12C with the protocol previously described. Then, 5 h post-transfection medium was changed and replaced with DMEM supplemented with 1% glutamine containing either 20 ng/mL DMOG, 20 ng/mL IL-6 or no stimulus (unstimulated). In addition, for cell transfected with KRASWT 15 ng/mL doxycycline was added for plasmid activation. Cell suspension samples were collected: once for all group during the seeding (24 h before transfection) then in triplicate for each group at 24 h, 48 h and 72 h post-transfection.

Medium samples were collected: once during the context introduction (5 h post-transfection) then in triplicate for each group at 24 h, 48 h and 72 h post-transfection. Cell suspension samples were used for cell counting using Scepter™ 2.0 Automated Cell Counter with 60 µm sensors (Merck Millipore), for cell viability and cellular ATP assessments using CellTiter-Glo® Luminescent Cell Viability Assay (Promega) and for Western-Blots of FLAG-KRAS (Sigma-Aldrich, Anti-FLAG® M2, F3165, mouse / monoclonal, 1:1000 dilution) normalized with β-actin (Cell Signaling, #4970, rabbit / monoclonal, 1:3000 dilution) using protocol previously described. For the Western Blot, after cell lysis, for each plasmid and at each time, the 3 context replicates were pooled to obtain enough protein to prepare 40 µL loading solution at 0.25 µg/µL (e.g., for WT at 24 h, the 3 replicates are 1 Unstim, 1 DMOG and 1 IL-6 samples). Medium were used for assessment of glucose uptake and lactate release using, respectively, Glucose-Glo™ and Lactate-Glo™ assays (Promega).

Statistical analysis

All data expressed as average ± standard deviations (SD), with SD represented by error bars. Statistical comparisons between two groups (typically treated group against control samples) were performed using a t-test. The average value and SD were calculated from at least three biological experiments. All tests were performed with a p-value of 0.05 using GraphPad Prism 9 software.

Network reconstruction and random walks analysis of AP-MS data

The starting point of the network are the 56 potential effectors of KRAS (*8*). Then, beginning from these effectors, STRING (version 11.5) was used to construct the network (*81*). All nodes which had a shortest path of 4 or less to these effectors were included, while filtering

out edges with a STRING confidence score of less than 0.7. For KRAS, only edges towards the effectors were included in the network. Apart from the KRAS-effector edges, all interactions in the network are considered undirected. The final network consists of 15,062 nodes and 493,838 edges.

Using the network, targeted random walks were performed starting from KRAS, in the following called source, for each target protein in each condition of interest. For a predetermined number of steps, based on the current node, one of the connected nodes is randomly chosen. For each random walk, there is always only one target protein. As soon as the target protein is reached, or a certain number of iterations has been exceeded, the walk ends. The random walks are biased towards proteins found in the interactome of a certain condition. This is facilitated by favoring nodes found in the interactome by a factor of 20 over nodes not found the interactome of the specific condition. The actual probability depends on the number of connecting nodes.

$$P(in\ APMS) = 20 * P(not\ in\ APMS)$$

The number of iterations for the random walk for each target is dynamically calculated based on the length of the shortest path between source and target.

$$walklen = shortestpath + 2$$

Finally, the number of random walks, limited by runtime and memory, was set to 100,000,000 for each target for each condition. The code for the random walks was written in python using numpy, scipy, pandas, numba and crsgraph. The script used to run this analysis is available on Zenodo (https://zenodo.org/record/6896565#.Yt6m63ZByUk) (DOI: 10.5281/zenodo.6896565).

Analysis of the random walks was performed by filtering out any paths that were found less than 100/100.000.000 walks and selecting the top 10 identified paths for each

target by frequency. Paths were decomposed into a sequence of edges, and all edges for one condition were concatenated to generate a condition-specific network of information flow from KRAS to all proteins associated with a specific GO term. Networks were visualized, and the effector layer of each network was extracted and visualized together.

All analysis and visualization were performed using R, in particular the packages dplyr, tidyr, stringr, purrr, furrr, ggplot and ggraph (*62-64*, *82*, *83*).

Supplementary Materials

Fig. S1. Classification of 56 effectors into groups based on their requirement of additional domains for efficient binding to Ras•GTP.

Fig. S2. Exogenous expression of FLAG-KRAS WT and mutant variants.

Fig. S3. Outlier analysis of AP-MS data.

Fig. S4. KRAS protein abundances across the whole AP-MS dataset.

Fig. S5. UMAP analysis performed on LFQ intensity and executed with MS log2-transformed data after filtering on the whole AP-MS dataset.

Fig. S6. Domain composition and capabilities of PM recruitment of 11 effectors identified in the AP-MS experiments.

Fig. S7. Principle component (PC) analysis after transforming the data set by summing up LFQ intensities for each GO biological process.

Fig. S8. Functional differences in the KRAS interactome.

Fig. S9. Visualization of the functional analysis of AP-MS experiments.

Fig. S10. Time course of exogenous expression of FLAG-KRAS WT and mutant during the Caco-2 phenotypical assessment experiment.

Fig. S11. Effect of culture and genetic contexts on phenotypic and metabolic parameters of Caco-2 cells.

Fig. S12. Examples for random walk analyses linked to GO term “Epithelial cell proliferation” (GO:0050673).

Fig. S13. Examples for random walk analyses linked to GO term “Regulation of glucose metabolic process” (GO:0010906).

Fig. S14. Examples for random walk analyses linked to GO term “ATP metabolic process” (GO:0046034).

Table S1. List of context-specific significant pairwise comparison of the sum of LFQ intensity of proteins in the AP-MS from six GO terms of interest.

Data file S1. Raw AP-MS LFQ intensity data with biological triplicates.

Data file S2. Log2 transformed AP-MS data with biological triplicates.

Acknowledgments

The authors would like to thank all members of the Kiel lab for discussions and critical reading of the manuscript. This work received funding from Science Foundation Ireland grant 16/FRL/3886 (CK) and from the Comprehensive Molecular Analytical Platform (CMAP) under The SFI Research Infrastructure Programme 18/RI/5702 (KW).

Author contributions

Methodology: CT, PJ, TS, SC, GO, KW, CK

Investigation: CT, PJ, TS, SC, GO, KW, CK

Writing – original draft: CT, TS, PJ, CK

Writing – review & editing: PJ, TS, SC, GO, KW

Disclosure and competing interests statement

Authors declare that they have no competing interests.

Data Availability

All data are available in the main text or the supplementary materials. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*54*) partner repository with the dataset identifier PXD035399. Data processing and analysis pipeline as well as results are available at Zenodo (https://zenodo.org/record/6896565#.Yt6m63ZByUk) (DOI: 10.5281/zenodo.6896565). Access to the AP-MS data through the Shiny app is available via GitHub (https://github.com/PhilippJunk/kras_apms_vis).

References

1. M. Vidal, M. E. Cusick, A. L. Barabási, Interactome networks and human disease. *Cell* **144**, 986-998 (2011).
2. A. H. van Boxel-Dezaire, M. R. Rani, G. R. Stark, Complex modulation of cell type-specific signaling in response to type I interferons. *Immunity* **25**, 361-372 (2006).

3. D. E. Hammond *et al.*, Differential reprogramming of isogenic colorectal cancer cells by distinct activating KRAS mutations. *J Proteome Res* **14**, 1535-1546 (2015).

4. C. Kiel, L. Serrano, Structure-energy-based predictions and network modelling of RASopathy and cancer missense mutations. *Mol Syst Biol* **10**, 727 (2014).

5. C. Kiel, D. Matallanas, W. Kolch, The Ins and Outs of RAS Effector Complexes. *Biomolecules* **11**, (2021).

6. S. A. Kennedy *et al.*, Extensive rewiring of the EGFR network in colorectal cancer cells expressing transforming levels of KRAS(G13D). *Nat Commun* **11**, 499 (2020).

7. D. K. Simanshu, D. V. Nissley, F. McCormick, RAS Proteins and Their Regulators in Human Disease. *Cell* **170**, 17-33 (2017).

8. V. Ibáñez Gaspar, S. Catozzi, C. Ternet, P. J. Luthert, C. Kiel, Analysis of Ras-effector interaction competition in large intestine and colorectal cancer context. *Small GTPases* **12**, 209-225 (2021).

9. C. Kiel, E. Verschueren, J. S. Yang, L. Serrano, Integration of protein abundance and structure data reveals competition in the ErbB signaling network. *Sci Signal* **6**, ra109 (2013).

10. R. C. Gimple, X. Wang, RAS: Striking at the Core of the Oncogenic Circuitry. *Front Oncol* **9**, 965 (2019).

11. J. M. Shields, K. Pruitt, A. McFall, A. Shaub, C. J. Der, Understanding Ras: 'it ain't over 'til it's over'. *Trends Cell Biol* **10**, 147-154 (2000).

12. S. Rezaei Adariani *et al.*, A comprehensive analysis of RAS-effector interactions reveals interaction hotspots and new binding partners. *J Biol Chem* **296**, 100626 (2021).

13. S. Wohlgemuth *et al.*, Recognizing and defining true Ras binding domains I: biochemical analysis. *J Mol Biol* **348**, 741-758 (2005).

14. C. Kiel *et al.*, Recognizing and defining true Ras binding domains II: in silico prediction based on homology modelling and energy calculations. *J Mol Biol* **348**, 759-775 (2005).

15. S. Catozzi, M. Halasz, C. Kiel, Predicted 'wiring landscape' of Ras-effector interactions in 29 human tissues. *NPJ Syst Biol Appl* **7**, 10 (2021).

16. E. Castellano, J. Downward, RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer* **2**, 261-274 (2011).

17. B. N. Kholodenko, J. B. Hoek, H. V. Westerhoff, Why cytoplasmic signalling proteins should be recruited to cell membranes. *Trends Cell Biol* **10**, 173-178 (2000).

18. R. L. Siegel *et al.*, Colorectal cancer statistics, 2020. *CA Cancer J Clin* **70**, 145-164 (2020).

19. E. R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-767 (1990).

20. G. A. Hobbs, C. J. Der, K. L. Rossman, RAS isoforms and mutations in cancer at a glance. *J Cell Sci* **129**, 1287-1292 (2016).

21. J. G. Tate *et al.*, COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47**, D941-d947 (2019).

22. D. Romano *et al.*, Protein interaction switches coordinate Raf-1 and MST2/Hippo signalling. *Nat Cell Biol* **16**, 673-684 (2014).

23. H. B. Engin, D. Carlin, D. Pratt, H. Carter, Modeling of RAS complexes supports roles in cancer for less studied partners. *BMC Biophys* **10**, 5 (2017).

24. J. Fogh, W. C. Wright, J. D. Loveless, Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J Natl Cancer Inst* **58**, 209-214 (1977).

25. M. J. Waldner, S. Foersch, M. F. Neurath, Interleukin-6--a key regulator of colorectal cancer development. *Int J Biol Sci* **8**, 1248-1253 (2012).

26. J. Zeng, Z. H. Tang, S. Liu, S. S. Guo, Clinicopathological significance of overexpression of interleukin-6 in colorectal cancer. *World J Gastroenterol* **23**, 1780-1786 (2017).

27. B. Ancrile, K. H. Lim, C. M. Counter, Oncogenic Ras-induced secretion of IL6 is required for tumorigenesis. *Genes Dev* **21**, 1714-1719 (2007).

28. H. Kikuchi, M. S. Pino, M. Zeng, S. Shirasawa, D. C. Chung, Oncogenic KRAS and BRAF differentially regulate hypoxia-inducible factor-1alpha and -2alpha in colon cancer. *Cancer Res* **69**, 8499-8506 (2009).

29. S. Y. Chun *et al.*, Oncogenic KRAS modulates mitochondrial metabolism in human colon cancer cells by inducing HIF-1α and HIF-2α target genes. *Mol Cancer* **9**, 293 (2010).

30. S. Zhang *et al.*, Stabilization of Hypoxia-inducible Factor by DMOG Inhibits Development of Chronic Hypoxia-Induced Right Ventricular Remodeling. *J Cardiovasc Pharmacol* **67**, 68-75 (2016).

31. H. H. Hsu *et al.*, Prostaglandin E2-Induced COX-2 Expressions via EP2 and EP4 Signaling Pathways in Human LoVo Colon Cancer Cells. *Int J Mol Sci* **18**, (2017).

32. N. Smakman *et al.*, Dual effect of Kras(D12) knockdown on tumorigenesis: increased immune-mediated tumor clearance and abrogation of tumor malignancy. *Oncogene* **24**, 8338-8342 (2005).

33. A. Greenhough *et al.*, The COX-2/PGE2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment. *Carcinogenesis* **30**, 377-386 (2009).

34. M. Y. Hein *et al.*, A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712-723 (2015).

35. A. L. Richards, M. Eckhardt, N. J. Krogan, Mass spectrometry-based protein-protein interaction networks for the study of human diseases. *Mol Syst Biol* **17**, e8792 (2021).

36. V. Beltran-Sastre *et al.*, Tuneable endogenous mammalian target complementation via multiplexed plasmid-based recombineering. *Sci Rep* **5**, 17432 (2015).

37. M. W. Dorrity, L. M. Saunders, C. Queitsch, S. Fields, C. Trapnell, Dimensionality reduction by UMAP to visualize physical and genetic interactions. *Nat Commun* **11**, 1537 (2020).

38. S. Catozzi *et al.*, Reconstruction and analysis of a large-scale binary Ras-effector signaling network. *Cell Commun Signal* **20**, 24 (2022).

39. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).

40. E. L. Huttlin *et al.*, Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* **184**, 3022-3040.e3028 (2021).

41. S. Eisenberg, Y. I. Henis, Interactions of Ras proteins with the plasma membrane and their roles in signaling. *Cell Signal* **20**, 31-39 (2008).

42. K. Miller-Jensen, K. A. Janes, J. S. Brugge, D. A. Lauffenburger, Common effector processing mediates cell-specific responses to stimuli. *Nature* **448**, 604-608 (2007).

43. M. Uhlén *et al.*, Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

44. E. L. Huttlin *et al.*, The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425-440 (2015).

45. E. L. Huttlin *et al.*, Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505-509 (2017).

46. M. H. Schaefer, L. Serrano, Cell type-specific properties and environment shape tissue specificity of cancer genes. *Sci Rep* **6**, 20707 (2016).

47. C. Nogales *et al.*, Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends Pharmacol Sci* **43**, 136-150 (2022).

48. C. S. Hughes *et al.*, Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68-85 (2019).

49. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906 (2007).

50. N. Bache *et al.*, A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics. *Mol Cell Proteomics* **17**, 2284-2296 (2018).

51. F. Meier *et al.*, Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol Cell Proteomics* **17**, 2534-2545 (2018).

52. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).

53. J. Cox *et al.*, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526 (2014).

54. Y. Perez-Riverol *et al.*, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-d552 (2022).

55. C. Lazar, L. Gatto, M. Ferro, C. Bruley, T. Burger, Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *J Proteome Res* **15**, 1116-1125 (2016).

56. L. Gatto, K. S. Lilley, MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288-289 (2012).

57. L. Gatto, S. Gibb, J. Rainer, MSnbase, Efficient and Elegant R-Based Processing and Visualization of Raw Mass Spectrometry Data. *J Proteome Res* **20**, 1063-1069 (2021).

58. J. Rainer *et al.*, A Modular and Expandable Ecosystem for Metabolomics Data Annotation in R. *Metabolites* **12**, (2022).

59. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289-300 (1995).

60. P. J. Luthert, C. Kiel, Combining Gene-Disease Associations with Single-Cell Gene Expression Data Provides Anatomy-Specific Subnetworks in Age-Related Macular Degeneration. *Netw Syst Med* **3**, 105-121 (2020).

61. A. P. Beckerman, D. Z. Childs, O. L. Petchey, in *Getting Started with R*. (2017), pp. 57-78.

62. H. Wickham *et al.*, Welcome to the Tidyverse. *Journal of Open Source Software* **4**, (2019).

63. H. Wickham, stringr: modern, consistent string processing. *The R Journal* **2**, (2010).

64. T. Mailund, in *R Data Science Quick Reference*. (2019), chap. Chapter 6, pp. 83-107.

65. X. Zhang *et al.*, Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nature Protocols* **13**, 530-550 (2018).

66. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

67. B. Phipson, S. Lee, I. J. Majewski, W. S. Alexander, G. K. Smyth, ROBUST HYPERPARAMETER ESTIMATION PROTECTS AGAINST HYPERVARIABLE GENES AND IMPROVES POWER TO DETECT DIFFERENTIAL EXPRESSION. *Ann Appl Stat* **10**, 946-963 (2016).

68. S. Oba *et al.*, A Bayesian missing value estimation method for gene expression profile data. *Bioinformatics* **19**, 2088-2096 (2003).

69. W. Stacklies, H. Redestig, M. Scholz, D. Walther, J. Selbig, pcaMethods--a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**, 1164-1167 (2007).

70. T. Wu *et al.*, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).

71. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**, 284-287 (2012).

72. M. Ashburner *et al.*, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29 (2000).

73. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* **49**, D325-d334 (2021).

74. S. Holm, A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* **6**, 65-70 (1979).

75. A. Schlicker, F. S. Domingues, J. Rahnenführer, T. Lengauer, A new measure for functional similarity of gene products based on Gene Ontology. *BMC Bioinformatics* **7**, 302 (2006).

76. G. Yu *et al.*, GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* **26**, 976-978 (2010).

77. Z. Gu, D. Hübschmann, Simplify enrichment: A bioconductor package for clustering and visualizing functional enrichment results. *Genomics Proteomics Bioinformatics*, (2022).

78. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).

79. C. Sievert, in *Interactive Web-Based Data Visualization with R, plotly, and shiny*. (2020), pp. 239-310.

80. Z. Gu, D. Hübschmann, Make Interactive Complex Heatmaps in R. *Bioinformatics* **38**, 1460-1462 (2021).

81. D. Szklarczyk *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).

82. T. L. Pedersen. An Implementation of Grammar of Graphics for Graphs and Networks [R package ggraph version 2.0.4]. (2020).

83. L. Wilkinson, ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics* **67**, 678-679 (2011).

Figure Legends

Fig. 1. Global analysis of context-specific KRAS WT and mutant interactomes. (**A**) Number of proteins identified in KRAS WT and mutant AP-MS experiments after filtering. The conditions are unstimulated (minimal medium), DMOG (20 and 200 ng/ml in minimal medium), EGF (20 and 200 ng/ml in minimal medium), IL6 (20 and 200 ng/ml in minimal medium), PGE2 (20 and 200 ng/ml in minimal medium), and TNFα (20 and 200 ng/ml in minimal medium). (**B**) Overlap of all interactions

identified in at least one condition with the literature KRAS interactome described in Kiel et al, Biomolecules, 2021. (**C**) Overlap of the same datasets as in panel B but focusing on effector proteins. (**D**) Principal component (PC) analysis performed on label-free quantification intensity (LFQi) and executed with MS log2- transformed data after filtering on the whole AP-MS dataset. Colors indicate the different growth conditions, i.e., DMOG, EGF, IL-6, PGE2, TNF-a and unstimulated (unstim), and shapes indicate the concentration of the conditions (none, 20 ng/ml and 200 ng/ml).

Fig. 2. Ras effector abundances, binding affinities, and detection in AP-MS experiments. Ras effector protein abundances in human colon tissue based on Wang et al., 2019 and Gaspar et al., 2021 (**A**) and mRNA abundances in Caco-2 cells based on the Human Protein Atlas database (**B**). The colors of the effectors' names at the bottom of each histogram correspond to Ras-effector binding affinities. The black star indicates effectors that were detected in at least one of the AP-MS experiments conducted in this work. TPM: transcripts per million; nTPM: normalised transcript expression values per sample; Kd: dissociation constant. (**C**) - (**E**) Effectors identified in at least one condition in the literature KRAS interactome described in Kiel et al, Biomolecules, 2021 (panel C), in this work (panel D), and in the combined datasets (panel E). Group 1, group 2 and group 3 effectors are normalized based on the total number of effectors in the respective group. (**F**) Number of conditions where an effectors is present.

Fig. 3. Summary of effector presence in AP-MS experiments in Caco-2 cells. The rows display the effectors in complex with KRAS grouped by mutational status (WT, G12D, G12V and G12C) and the columns represent the conditions (unstimulated or stimulated with either DMOG, EGF, IL-6, TNF-a and PGE2). To generate the two heatmaps, the LFQ intensities were analyzed, the different concentrations for each

stimulation merged and translate for the heatmap (**A**) in terms of presence or absence of an effector (color code: detected = blue and not detected = yellow) and for second heatmap (**B**) the LFQ intensities were directly plotted into the heatmap (code: from low abundance = yellow to high abundance = blue). The heatmaps were created using GraphPad Prism 9.

Fig. 4. Gene set enrichment analysis (GSEA) of AP-MS analysis in different genetic and culture contexts. (**A**) Overview of the functional analysis pipeline. The first approach is a standard differential analysis pipeline using limma, followed by a gene set enrichment analysis against the "Biological Process" GO ontology. The second approach is to sum up LFQ intensities for each of the "Biological Process" GO ontology terms. Then, for each term, an ANOVA was performed to identify whether there was a significant influence of mutation status, condition and concentration and their interaction terms. Both analysis were then analyzed using semantic analysis of ontology terms. (**B**). Distance matrix with the pairwise semantic distances between the GO terms is shown in the center of the plot. On the right, the bigger clusters are annotated by word clouds. On the left, the data from the ANOVA and the GSEA analysis pipeline is shown, for the ANOVA whether a main effect is statistically significant and for the GSEA, whether a significant enrichment was found in the relevant pairwise comparison. All data shown is in comparison to the WT and unstimulated for the analysis of mutation status and condition, respectively.

Fig. 5. Comparison of phenotypic parameters with the sum of LFQ intensity of related GO terms in the AP-MS. (**A**) Radar plot of phenotypic parameters related to cell proliferation, glucose metabolism and ATP metabolism of Caco-2 cells under different genetic and culture contexts. (**B**) Radar plot of the GO terms related to cell proliferation, glucose metabolism and ATP metabolism of Caco2 cells under different

'culture' and 'genetic' contexts. Value in radar plots are the LS mean values normalized by the maximum value of each parameter to obtain values between 0 and 1. (**C**) Pearson correlation matrix of phenotypic parameters versus sum of LFQ intensity of specific GO terms from the AP-MS. Phenotypic parameters and GO terms related to the same biological process are displayed with the same colors.

Fig. 6. Pathway analysis by biased random walks. (**A**) Heatmaps of effector traversal in the different genetic and culture for selected GO terms associated with cell phenotypes. Epithelial cell proliferation (GO:0050673). (**B**) Positive regulation of cell population proliferation (GO:0008284). (**C**) Glycolytic process (GO:0006096). (**D**) Regulation of glucose metabolic process (GO:0010906). (**E**) ATP metabolic process (GO:0046034). (**F**) regulation of ATP metabolic process (GO:1903578). The colour scale indicates how often an effector is traversed.

Figure 1

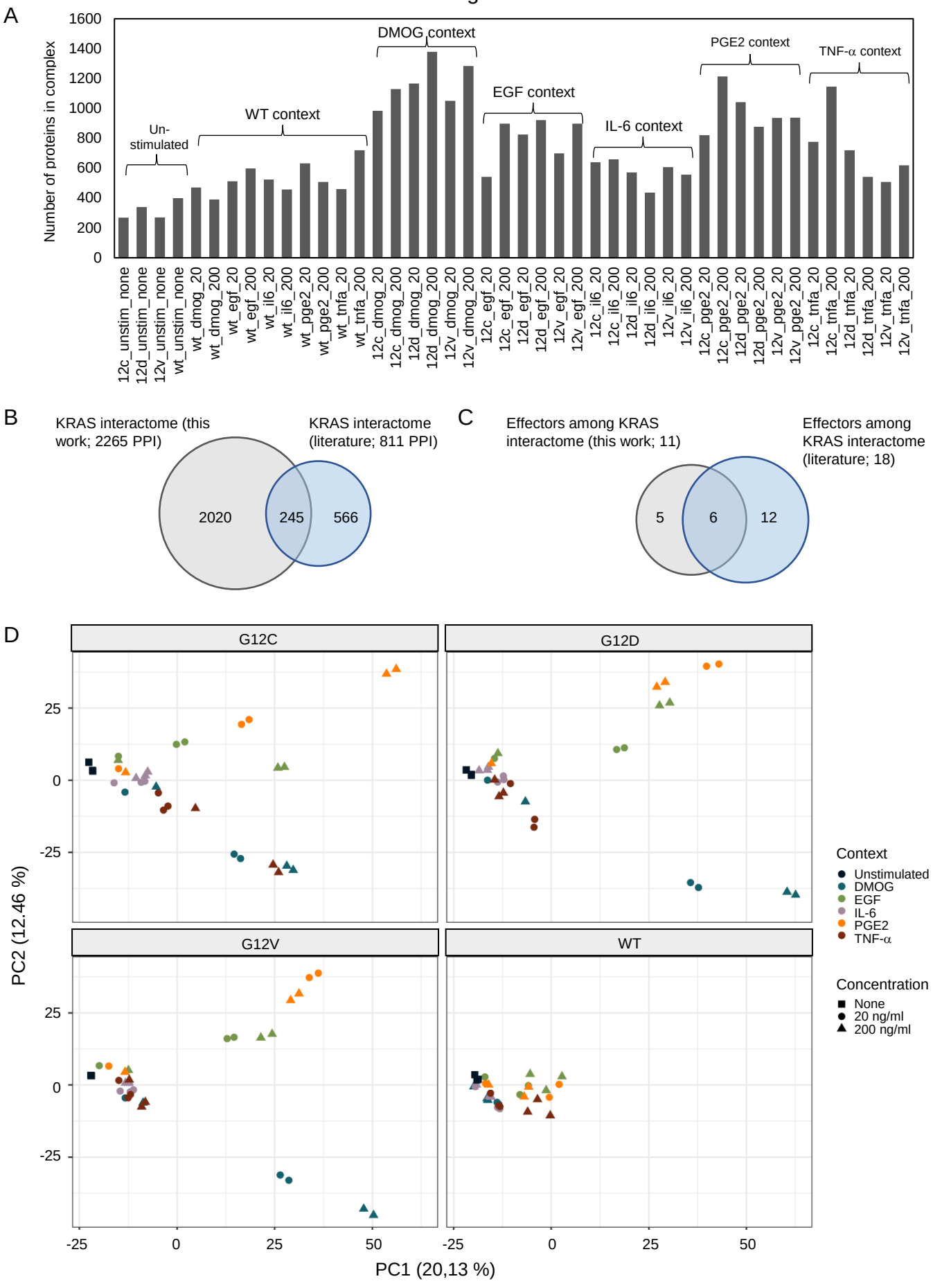
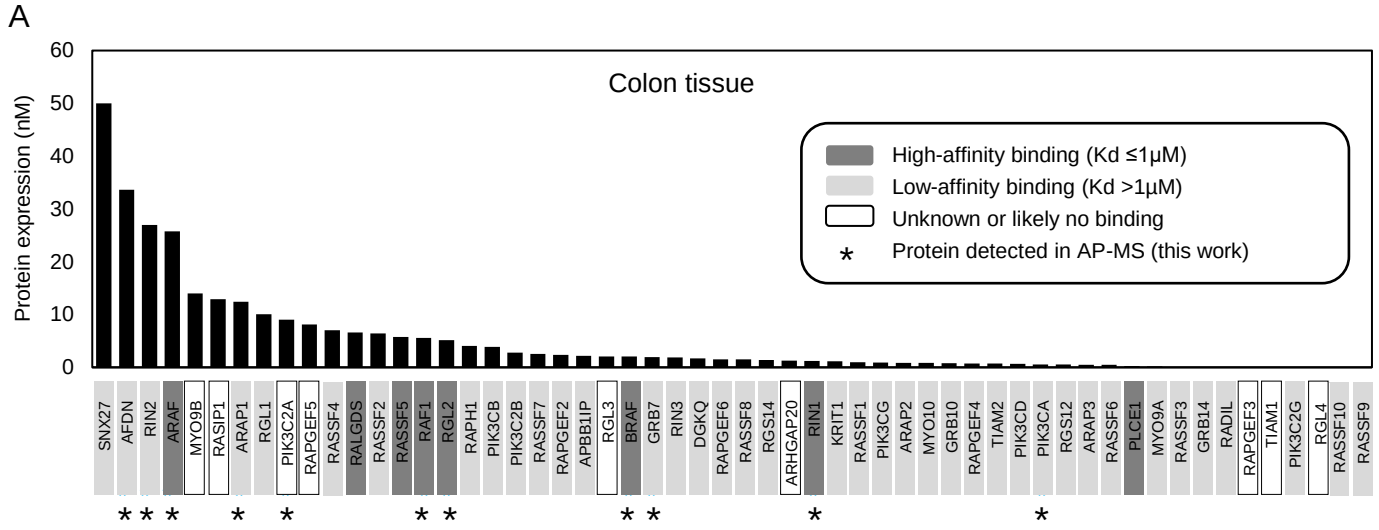
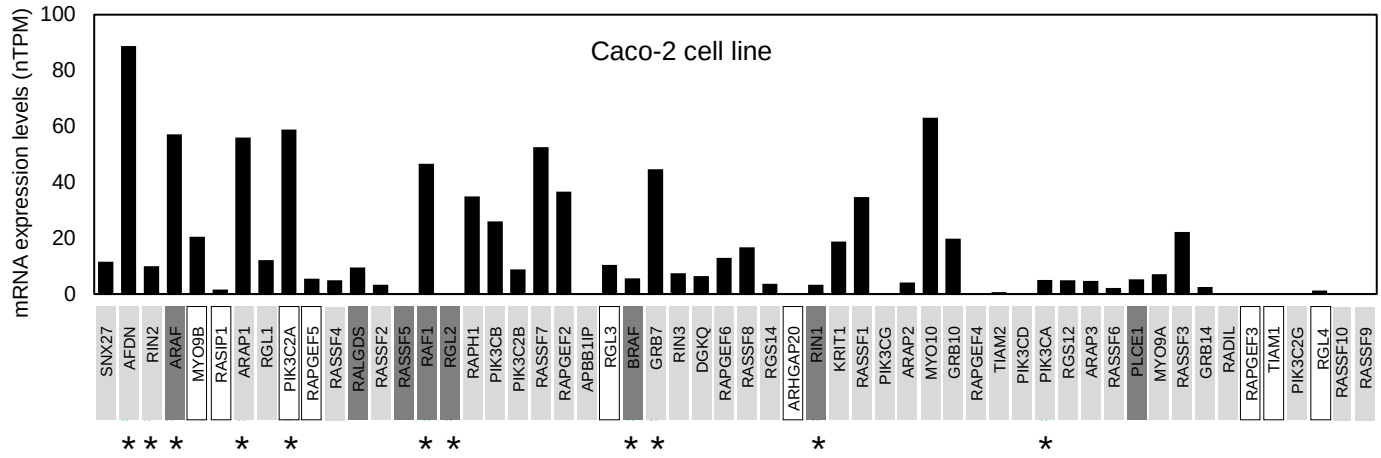
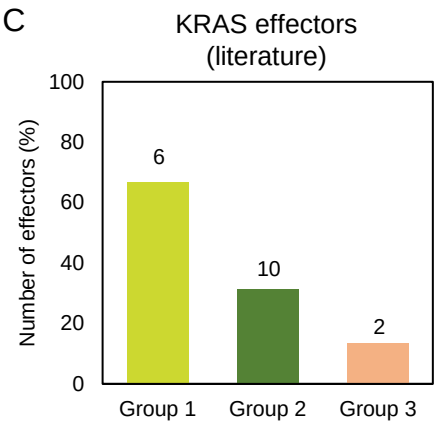
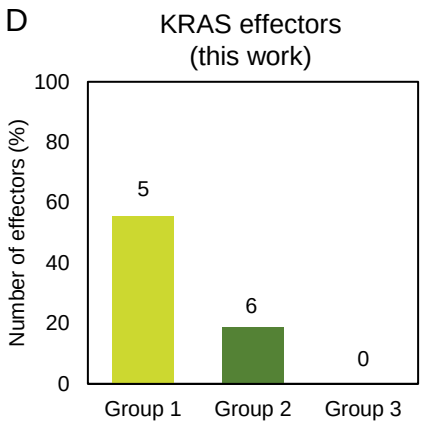

A
Number of proteins in complex
1600
1400
1200
1000
800
600
400
200
0
Un-stimulated
WT context
DMOG context
EGF context
IL-6 context
PGE2 context
TNF-α context
12c_unstim_none
12d_unstim_none
12v_unstim_none
wt_unstim_none
wt_dmog_20
wt_dmog_200
wt_egf_20
wt_egf_200
wt_il6_20
wt_il6_200
wt_pge2_20
wt_pge2_200
wt_tnfa_20
wt_tnfa_200
12c_dmog_20
12c_dmog_200
12d_dmog_20
12d_dmog_200
12v_dmog_20
12v_dmog_200
12c_egf_20
12c_egf_200
12d_egf_20
12d_egf_200
12v_egf_20
12v_egf_200
12c_il6_20
12c_il6_200
12d_il6_20
12d_il6_200
12v_il6_20
12v_il6_200
12c_pge2_20
12c_pge2_200
12d_pge2_20
12d_pge2_200
12v_pge2_20
12v_pge2_200
12c_tnfa_20
12c_tnfa_200
12d_tnfa_20
12d_tnfa_200
12v_tnfa_20
12v_tnfa_200
B
KRAS interactome (this work; 2265 PPI)
KRAS interactome (literature; 811 PPI)
2020
245
566
C
Effectors among KRAS interactome (this work; 11)
Effectors among KRAS interactome (literature; 18)
5
6
12
D
G12C
G12D
G12V
WT
PC2 (12.46 %)
PC1 (20,13 %)
25
0
-25
50
Context
Unstimulated
DMOG
EGF
IL-6
PGE2
TNF-α
Concentration
None
20 ng/ml
200 ng/ml

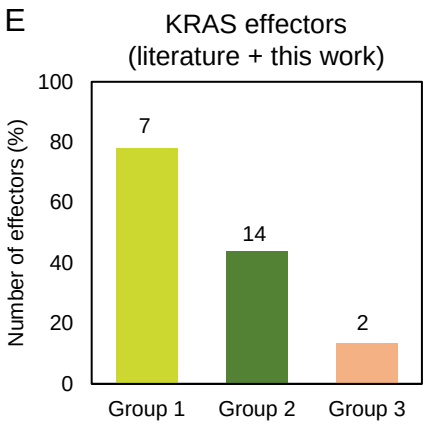
Figure 2


A
Colon tissue
Protein expression (nM)
High-affinity binding (Kd ≤1μM)
Low-affinity binding (Kd >1μM)
Unknown or likely no binding
Protein detected in AP-MS (this work)
SNX27 AFDN RIN2 ARAF MYO9B RASIP1 ARAP1 RGL1 PIK3C2A RAPGEF5 RASSF4 RALGDS RASSF2 RASSF5 RAF1 RGL2 RAPH1 PIK3CB PIK3C2B RASSF7 RAPGEF2 APBB1IP RGL3 BRAF GRB7 RIN3 DGKQ RAPGEF6 RASSF8 RGS14 ARHGAP20 RIN1 KRIT1 RASSF1 PIK3CG ARAP2 MYO10 GRB10 RAPGEF4 TIAM2 PIK3CD PIK3CA RGS12 ARAP3 RASSF6 PLCE1 MYO9A RASSF3 GRB14 RADIL RAPGEF3 TIAM1 PIK3C2G RGL4 RASSF10 RASSF9


B
Caco-2 cell line
mRNA expression levels (nTPM)
SNX27 AFDN RIN2 ARAF MYO9B RASIP1 ARAP1 RGL1 PIK3C2A RAPGEF5 RASSF4 RALGDS RASSF2 RASSF5 RAF1 RGL2 RAPH1 PIK3CB PIK3C2B RASSF7 RAPGEF2 APBB1IP RGL3 BRAF GRB7 RIN3 DGKQ RAPGEF6 RASSF8 RGS14 ARHGAP20 RIN1 KRIT1 RASSF1 PIK3CG ARAP2 MYO10 GRB10 RAPGEF4 TIAM2 PIK3CD PIK3CA RGS12 ARAP3 RASSF6 PLCE1 MYO9A RASSF3 GRB14 RADIL RAPGEF3 TIAM1 PIK3C2G RGL4 RASSF10 RASSF9


C
KRAS effectors (literature)
Number of effectors (%)
6
10
2
Group 1
Group 2
Group 3


D
KRAS effectors (this work)
Number of effectors (%)
5
6
0
Group 1
Group 2
Group 3


E
KRAS effectors (literature + this work)
Number of effectors (%)
7
14
2
Group 1
Group 2
Group 3

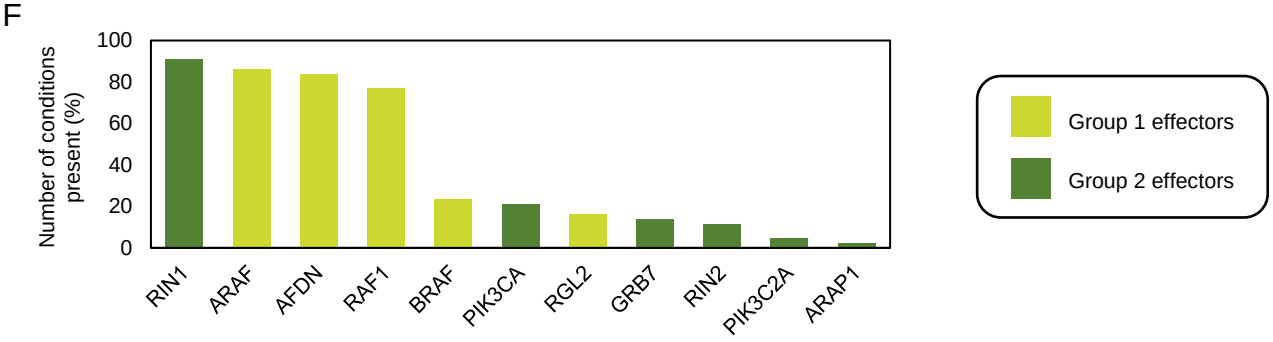

F
Number of conditions present (%)
RIN1 ARAF AFDN RAF1 BRAF PIK3CA RGL2 GRB7 RIN2 PIK3C2A ARAP1
Group 1 effectors
Group 2 effectors

Figure 3

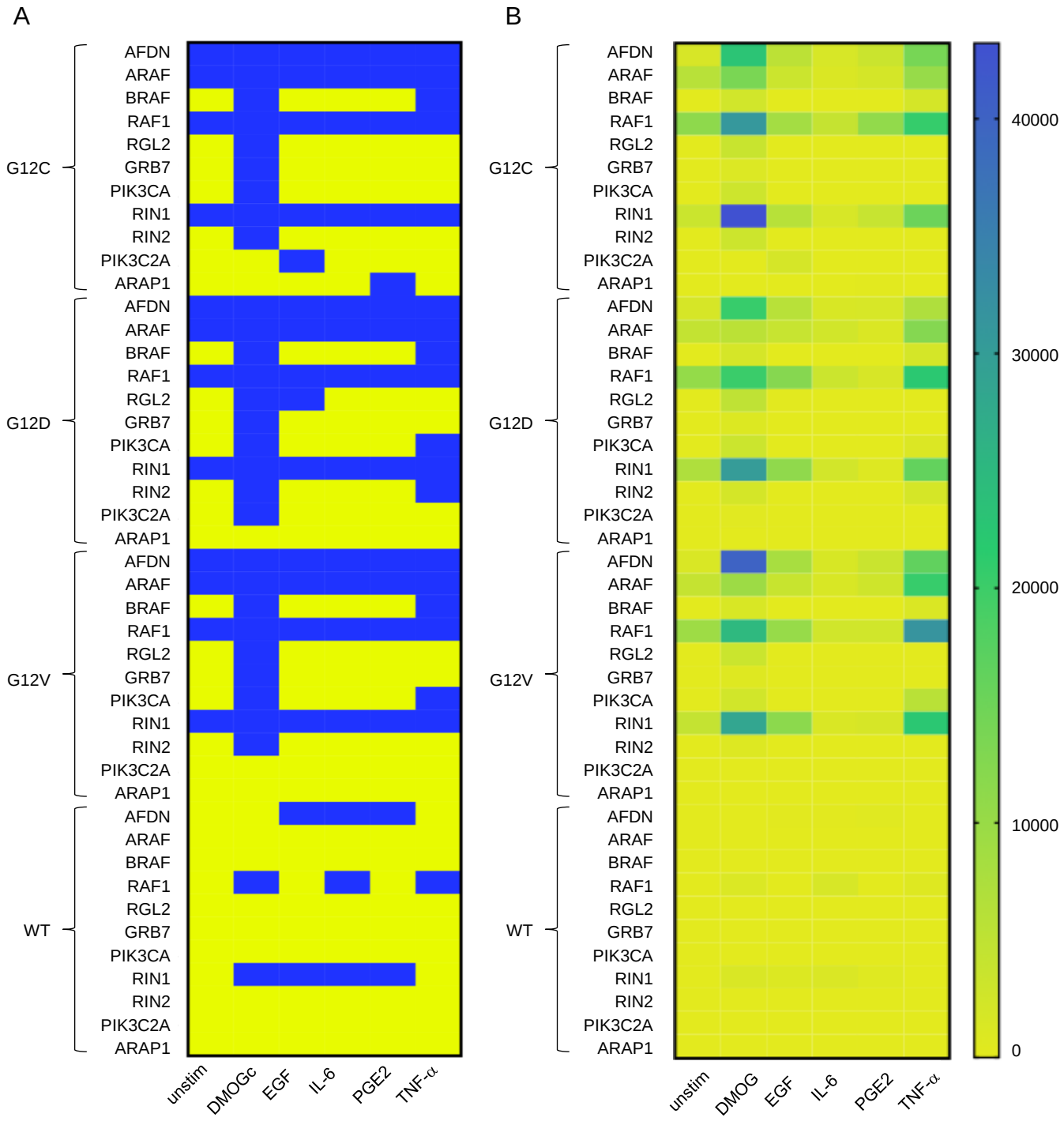
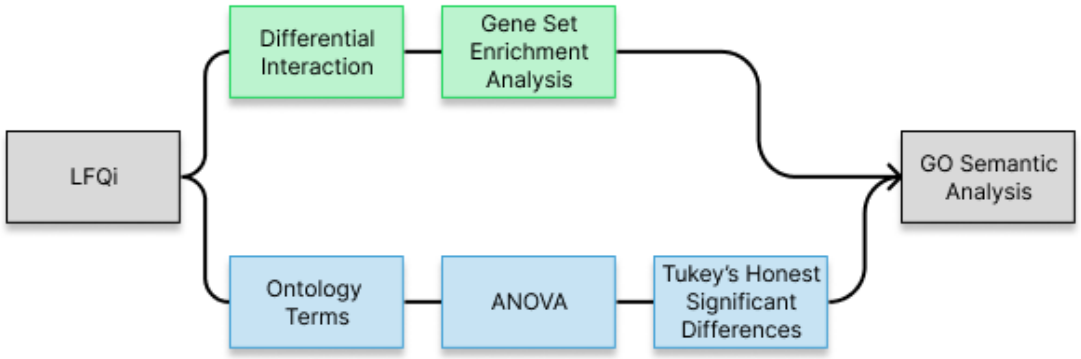
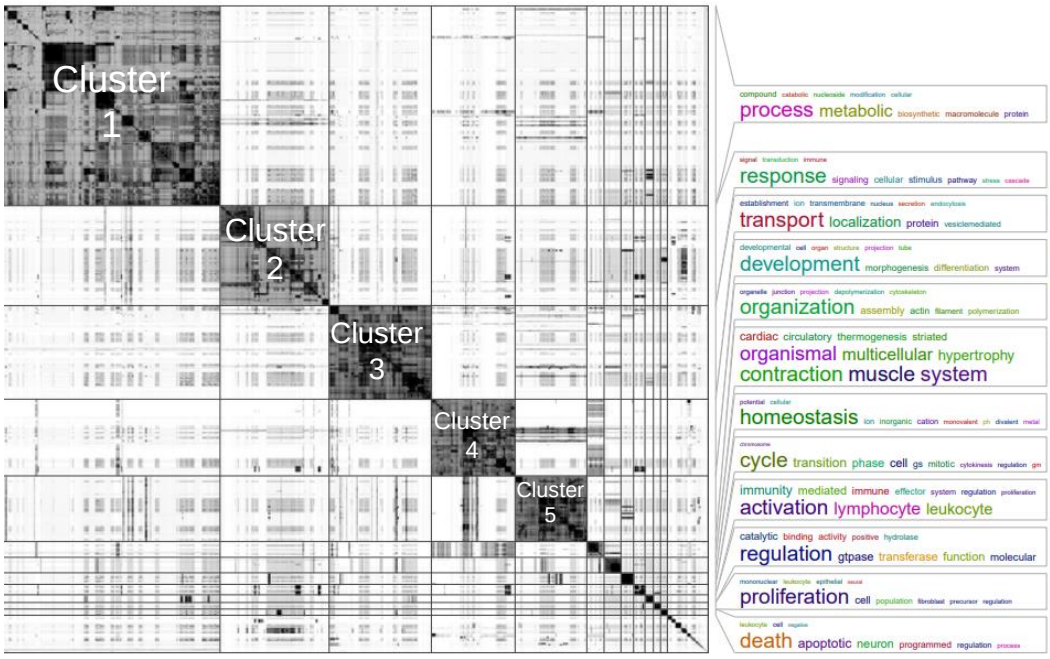

A
B
G12C
G12D
G12V
WT
AFDN
ARAF
BRAF
RAF1
RGL2
GRB7
PIK3CA
RIN1
RIN2
PIK3C2A
ARAP1
unstim
DMOG
EGF
IL-6
PGE2
TNF-α
40000
30000
20000
10000
0

Figure 4

A


Differential Interaction
Gene Set Enrichment Analysis
LFQi
GO Semantic Analysis
Ontology Terms
ANOVA
Tukey's Honest Significant Differences

B


Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
process metabolic
response signaling cellular stimulus pathway
transport localization protein
development morphogenesis differentiation
organization assembly actin
organismal multicellular hypertrophy
contraction muscle system
homeostasis
cycle transition phase cell
immunity mediated immune
activation lymphocyte leukocyte
regulation gtpase transferase function molecular
proliferation cell
death apoptotic neuron programmed regulation

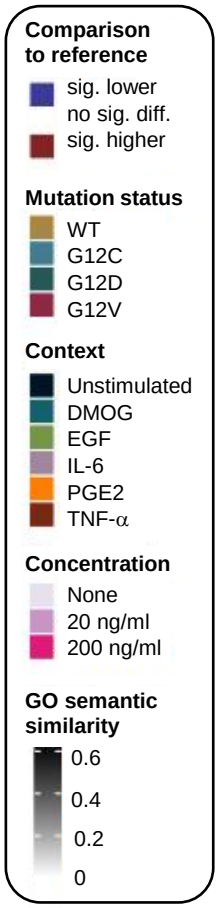

Comparison to reference
sig. lower
no sig. diff.
sig. higher
Mutation status
WT
G12C
G12D
G12V
Context
Unstimulated
DMOG
EGF
IL-6
PGE2
TNF-α
Concentration
None
20 ng/ml
200 ng/ml
GO semantic similarity
0.6
0.4
0.2
0

Figure 5

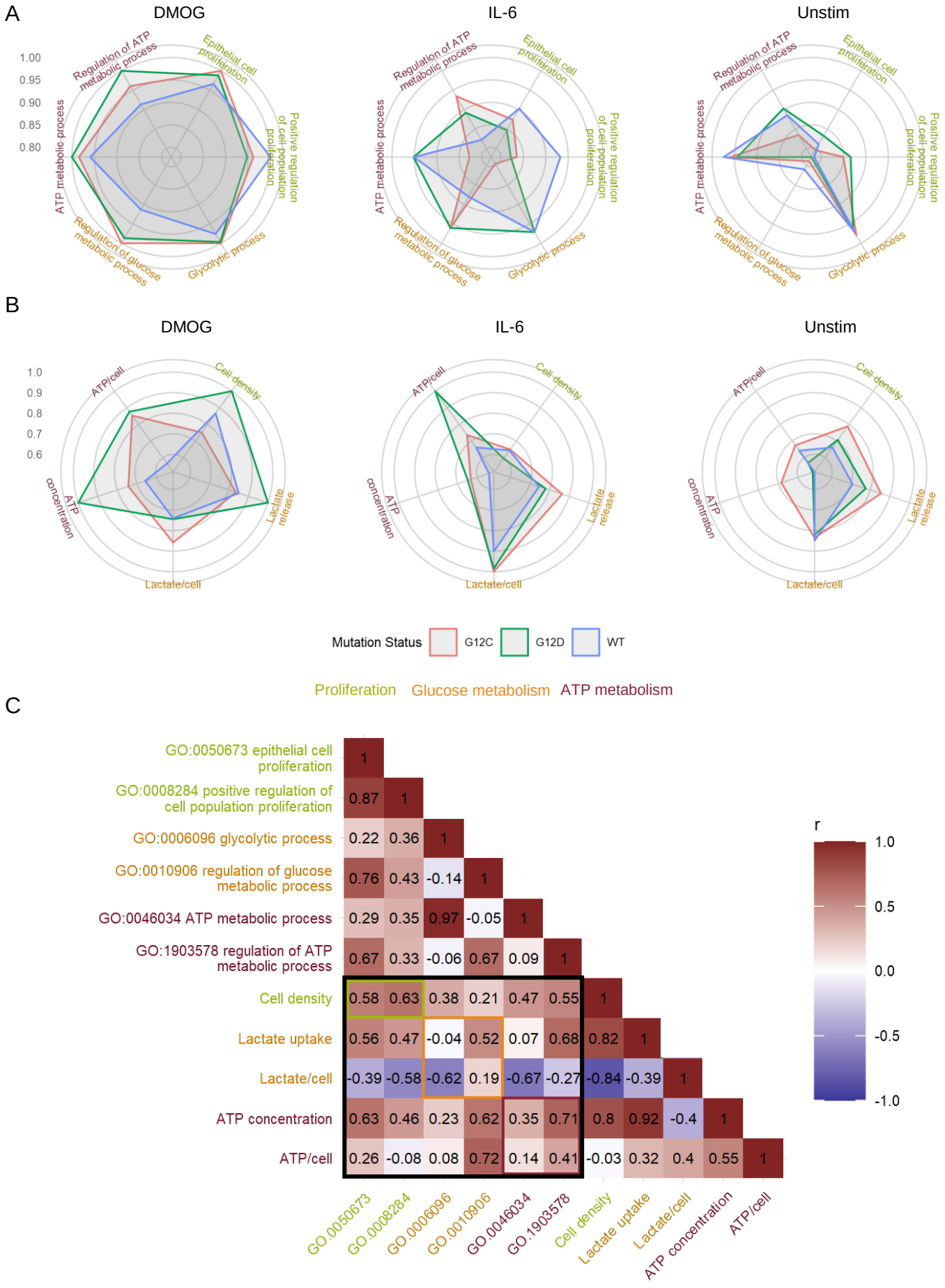
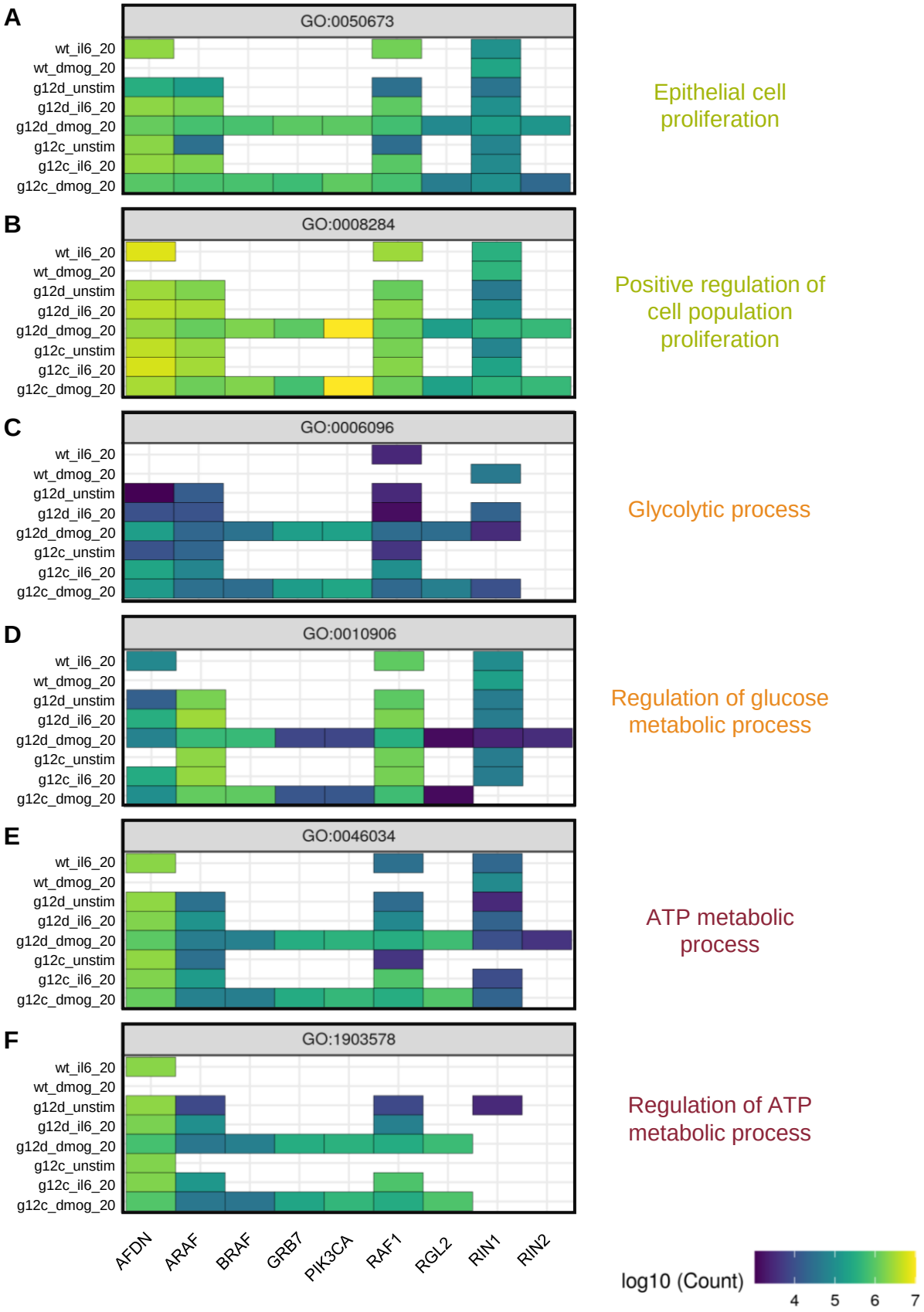

A
DMOG
IL-6
Unstim
1.00
0.95
0.90
0.85
0.80
Regulation of ATP metabolic process
Epithelial cell proliferation
Positive regulation of cell population proliferation
Glycolytic process
Regulation of glucose metabolic process
ATP metabolic process
B
DMOG
IL-6
Unstim
1.0
0.9
0.8
0.7
0.6
ATP/cell
Cell density
Lactate release
Lactate/cell
ATP concentration
Mutation Status
G12C
G12D
WT
Proliferation
Glucose metabolism
ATP metabolism
C
GO:0050673 epithelial cell proliferation
GO:0008284 positive regulation of cell population proliferation
GO:0006096 glycolytic process
GO:0010906 regulation of glucose metabolic process
GO:0046034 ATP metabolic process
GO:1903578 regulation of ATP metabolic process
Cell density
Lactate uptake
Lactate/cell
ATP concentration
ATP/cell
1
0.87 1
0.22 0.36 1
0.76 0.43 -0.14 1
0.29 0.35 0.97 -0.05 1
0.67 0.33 -0.06 0.67 0.09 1
0.58 0.63 0.38 0.21 0.47 0.55 1
0.56 0.47 -0.04 0.52 0.07 0.68 0.82 1
-0.39 -0.58 -0.62 0.19 -0.67 -0.27 -0.84 -0.39 1
0.63 0.46 0.23 0.62 0.35 0.71 0.8 0.92 -0.4 1
0.26 -0.08 0.08 0.72 0.14 0.41 -0.03 0.32 0.4 0.55 1
r
1.0
0.5
0.0
-0.5
-1.0

Figure 6


A
GO:0050673
B
GO:0008284
C
GO:0006096
D
GO:0010906
E
GO:0046034
F
GO:1903578
wt_il6_20
wt_dmog_20
g12d_unstim
g12d_il6_20
g12d_dmog_20
g12c_unstim
g12c_il6_20
g12c_dmog_20
AFDN
ARAF
BRAF
GRB7
PIK3CA
RAF1
RGL2
RIN1
RIN2
Epithelial cell proliferation
Positive regulation of cell population proliferation
Glycolytic process
Regulation of glucose metabolic process
ATP metabolic process
Regulation of ATP metabolic process
log10 (Count)
4
5
6
7