

Anterior Gradient proteins in gastrointestinal cancers: from cell biology to pathophysiology

Emeric Boisteau^{1,2}, Céline Posseme^{1,3,§}, Federico Di Modugno^{1,3,§}, Julien Edeline^{1,3},
Cédric Coulouarn¹, Roman Hrstka⁴, Andrea Martisova^{4,5}, Frédéric Delom⁶, Xavier
Treton⁷, Leif A. Eriksson⁸, Eric Chevet^{1,3*}, Astrid Lièvre^{1,2*}, Eric Ogier-Denis^{1,3*}

¹INSERM U1242, University of Rennes, Rennes, France. ²Department of Gastroenterology, University Hospital Pontchaillou, University of Rennes, Rennes, France. ³Centre de Lutte contre le cancer Eugène Marquis, Rennes, France. ⁴Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Brno, Czech Republic. ⁵National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic. ⁶University of Bordeaux, INSERM U1312, Bordeaux, France; ⁷Assistance Publique-Hôpitaux de Paris, University of Paris, Clichy, France; ⁸Department of Chemistry and Molecular Biology, University of Gothenburg, Göteborg, Sweden

[§] equal contribution

*Correspondence to EC (eric.chevet@inserm.fr) ORCID Id 0000-0001-5855-4522, AL (astrid.lievre@chu-rennes.fr) ORCID Id 0000-0003-2118-717X, and EOD (eric.ogier-denis@inserm.fr) ORCID Id 0000-0002-0057-7593

Abstract

Most of the organs of the digestive tract comprise secretory epithelia which require specialized molecular machines to achieve their functions. As such Anterior GRadient (AGR) proteins, which comprise AGR1, AGR2 and AGR3 belong to the Protein Disulfide Isomerase family, and are involved in secretory and transmembrane protein biogenesis in the endoplasmic reticulum (ER). They are generally expressed in epithelial cells with high levels in most of the digestive tract epithelia. To date, the vast majority of the reports concerns AGR2, which has been shown to exhibit various subcellular localizations and to exert pro-oncogenic functions. AGR2 overexpression has recently been associated to a poor prognosis in digestive cancers. AGR2 is also involved in epithelial homeostasis. Its deletion in mice results in severe diffuse gut inflammation, whereas in inflammatory bowel diseases, the secretion of AGR2 in the extracellular milieu participates to the reshaping of cellular microenvironment. AGR2 thus plays a key role in inflammation and oncogenesis and may represent a therapeutic target of interest. In this review, we summarize the already known roles and mechanisms of action of the AGR family proteins in digestive diseases, their expression in the healthy digestive tract and in digestive oncology. At last, we discuss the potential diagnostic and therapeutic implications underlying the biology of AGR proteins.

1. Introduction

The digestive system is composed of organs with specific differentiated cells, each of which having specific functions in the digestion process. Beyond these differences, all these organs show common secretory properties essential for homeostasis. Alterations in the secretory pathway have been implicated in oncogenesis. Digestive cancers, like most cancers, often result from a chronic inflammation that dysregulates the cell metabolism, gene expression and epigenetics. Chronic inflammation can be responsible for precancerous lesions, before the evolution toward invasive cancer, like in Barrett metaplasia in the esophagus, inflammatory bowel diseases (IBD) in the gut or pancreatic intraepithelial neoplasia (PanIN) in the pancreas. Among the cellular dysregulations that favor carcinogenesis, alterations of endoplasmic reticulum (ER) homeostasis and the subsequent adaptive mechanisms activated to cope with it were shown to significantly contribute to these processes. The ER is the first compartment of the secretory pathway that enables the proper biogenesis, folding, and maturation of secretory and transmembrane proteins (1). Beyond its role in protein homeostasis, the ER also maintains calcium and lipid homeostasis. As part of the ER folding and quality control machinery, the Protein Disulfide Isomerase (PDI) family is involved in the formation of inter and intra-molecular disulfide bridges in client proteins, productive protein folding, redox activity, and disulfide exchange through their thioredoxin-like domain (1-3). Although the roles of most PDI family members in protein folding have been well characterized, their function in ER homeostasis control and proteostasis maintenance is still unclear (4). Along with glutaredoxins and thioredoxins, PDI are members of the thioredoxin (TRX) superfamily which are characterized by the presence of a thioredoxin fold (5). So far, 20 PDIs have been identified (4,6). Most of them are located in the lumen of the ER, while others are transmembrane proteins

(TMX 1 to TMX 5) (6). PDI are defined as “proteins that can react with cysteine side chains”, but it includes also “proteins that contain non-thiol-reactive thioredoxin-like domains with chaperone-like activities for ER folding and secretion of proteins” (7). Therefore, the PDIs can be divided between thiol-non reactive (ERp27 and ERp29) and thiol-reactive proteins.

The AGR family is composed of three PDI-like proteins: AGR1 (also known as Thioredoxin domain-containing protein 12 (TXNDC12)), AGR2 (protein disulfide isomerase family A member 17 (PDIA17)), and AGR3 (Protein disulfide isomerase family A, member 18 (PDIA18)) (7,8). AGR1 is expressed in the ER and plays a role in the formation of mixed disulfide bounds with proteins (9). The expression of AGR2 is induced upon disruption of ER homeostasis, through the activation of the ATF6 and IRE1 arms of the unfolded protein response (UPR) (1). AGR2 is shown to contribute to protein folding and maturation (10,11) and to participate in the elimination of misfolded proteins through ER-associated degradation (ERAD) (12). Several studies in the last decade have shown that AGR2 expression is associated with poor prognosis in solid cancers and may play pro-oncogenic roles through yet uncharacterized mechanisms. AGR3 is also an ER resident protein involved in the formation of disulfide bonds but its expression is independent of ER stress (13). Compared to AGR2, less is known about the functions and regulations of both AGR1 and AGR3 as well as their respective roles in the digestive tract. In this review, we cover the most recent data on the AGR family from biochemical and cell biology perspectives; we next connect these properties to the functions of these proteins in both normal and pathological gastrointestinal tract and lastly discuss the potential theranostic aspects related to the AGR family.

2. The AGR family

2.1. Genes, structures and regulators

AGR2 is the most documented member of the AGR family. It was initially identified in the anterior region of the ectoderm of *Xenopus laevis* as XAG-2 (14–16). XAG-2 plays important roles in the endodermal patterning and in forebrain maturation by anteroposterior migration of the cement gland (mucus-secreting organs) during embryogenesis (14,17). In human, AGR2 is also expressed during embryonic development in the endoderm derived tissues like esophagus, trachea, lung, liver, stomach, and in the gut (18). In adult human tissues, the highest levels of AGR2 expression are observed in the gastrointestinal tract from the stomach to the rectum, in the genitourinary tract (urinary bladder), and in the respiratory system (see Figure 3). Human AGR2 is located on chromosome 7 (19) and is composed of 8 exons and 7 introns. Seven different AGR2 transcripts are reported in the Ensembl database. In cancer, the most recurrent AGR2 anomaly is amplification; only few mutations or deletions were described (20). AGR2 predominant isoform is 175 amino acid long resulting in a molecular weight of ~20 kDa. Three other isoforms have been described [B5MC07, C9J3E2 and H7C3Z9 (UniProt)] but so far, there is no information on their functions. Its structure is composed of an N-terminal cleavable signal peptide, a non-canonical thioredoxin domain CXXS, a peptide binding loop and a C-terminal domain ER-retention (KTEL) motif (**Figure 1a, b**) (1). AGR1 has a high similarity with the thioredoxin domain CXXC motif in invertebrates, but is lost in higher vertebrates where it is replaced by the CXXC containing protein TXNDC12 (21–25). It is a 18kDa ER resident protein composed of 146 amino-acid in its mature sequence (Human Protein Atlas - (26)). AGR1 gene is located on chromosome 1 unlike AGR2 and AGR3. Similar to other PDI, it is involved in disulfide bond formation (24). AGR2 and AGR3 are

present in both amphibians and higher vertebrates (8). AGR proteins ER retention motifs differ (i.e. EDEL for AGR1, KTEL for AGR2, QSEL for AGR3) as do their thioredoxin-like motifs (**Figure 1a, b**) (27). Even though both proteins are very similar (>90% identity) and highly conserved in mammals (e.g. mouse and human – **Figure 1b**), the role of AGR3 is less documented than AGR2 but its expression seems correlated with AGR2 and plays similar roles in some aspects, for instance promoting migration and metastasis (28). However, expression of AGR3 is independent of ER stress suggesting that the regulation of the two proteins differs each of another. AGR3 gene is located on chromosome 7p21.1, it is composed of 10 exons and can produce three different transcripts (AGR3-201, AGR3-202, AGR3-203). The molecular weight of AGR3-201 is 19.2 kDa and contains 166 amino acids (Human Protein Atlas – (29)). AGR3 was first identified in breast cancer cell membrane and in breast tumors, where its expression correlates with estrogen receptor expression (30,31). It is expressed in airway epithelium where it is required for muco-ciliary clearance and for regulation of ciliary beat frequency (32). It is also expressed in prostate, as well as stomach and liver tissues of the digestive tract (33).

AGR2 exists in monomeric or homodimeric forms (34). AGR2 dimerizes through E60 and K64-dependent salt bridges, far from the thioredoxin domain (35,36) (**Figure 1c**). It was also reported that AGR2 could form cysteine 81-dependent dimers but the precise functions of these dimers remain elusive (37). AGR2 was also found to form hetero dimers with AGR3 *in vitro*, but again the functions associated with this complex are unknown (38) (**Figure 1d**). AGR2 thioredoxin-like domain is found at position 81-84 (CXXS) and is conserved in mouse and human. It may be involved in the formation of mixed disulfide bridges with client proteins (39). ER stress conditions were shown to induce AGR2 monomerization, which was correlated with its secretion to the

extracellular milieu (34,40). Interestingly, AGR3 thioredoxin-like domain is not conserved in human and mouse, exhibiting a CXXS sequence in human and a canonical CXXC site in mouse (**Figure 1b**). This might reflect a difference in AGR3 redox functions between the species. The AGR2 peptide binding loop (VDPSL, conserved in AGR3, **Figure 1b**) interacts with proteins such as the AAA+ protein RuvBL2/Reptin (41). The AGR2 C-terminal domain contains a non-canonical ER retention motif (KTEL), that enables a partial retention of AGR2 in the ER through KDEL receptors (27).

2.2. AGR proteins functions and subcellular locations

AGR proteins exhibit various sub-cellular locations, associated with different roles (**Figure 2**). So far most if not all functions of AGR1 were associated with the ER localization (24) and in particular in the regulation of the ATF6 arm of the UPR (42,43). Even though AGR2 and AGR3 show all the features of ER-resident luminal proteins (*i.e.* signal peptide and ER-retention motif), both proteins are released in the extracellular environment and AGR2 was also found in the cytosol (44). In the ER, AGR2 (erAGR2) participates in protein folding, maturation and secretion. It is also involved in ER-associated protein degradation (1). In particular, erAGR2 contributes to the productive folding of i) mucins (MUC2, MUC5A/C) most likely through the formation of mixed disulfide bonds (13), ii) low affinity, Ca²⁺-binding, multiple EF-hand proteins such as CALU and RCN1 (45–48), iii) proteases such as Cathepsins B and D (45), iv) adhesion molecules such as EpCAM (49) or v) other PDIs (50). In contrast to AGR2, the ER functions of AGR3 remain to be fully characterized. The contribution of AGR2 to the productive folding of proteins whose functions are associated with cancer development, thereby sustaining their functions upon overexpression might represent

a pro-oncogenic feature (**Table 1**). In addition to the “canonical” ER localization of AGR2 and AGR3, both proteins were also found to be secreted extracellularly (eAGR2, eAGR3). Although the underlying mechanisms have not been studied for AGR3, it appears that ER stress is a driver of AGR2 secretion. Indeed, we demonstrated that upon ER stress induction either overexpression of AGR2 or dimer disruption (or both) correlated with enhanced AGR2 secretion (34). AGR2 dimer regulation was found to depend on the p24 cargo receptor family member TMED2 (34). It was also shown that the Cys81Ser mutation in the CXXS domain also prompted AGR2 secretion (51). Moreover, AGR2 KTEL motif binds to KDEL-R in a non-canonical manner, enabling an easier secretion of AGR2 compared to other proteins KDEL-containing motifs (1,52,53). This was confirmed upon replacement of KTEL by KDEL in AGR2 which led to its ER sequestration (52). The extracellular AGR2 (eAGR2) as well as eAGR3 plays multiple pathophysiological roles in inflammation, cell proliferation, angiogenesis, cell migration and metastasis, through autocrine, paracrine and maybe endocrine secretion (**Table 1**). The third sub-cellular localization in which AGR2 was detected is the cytosol. The mechanism by which AGR2 might relocate from the ER lumen to the cytosol was identified to be an ER stress-induced reflux of properly folded proteins initially identified in yeast and also found in mammalian cells (44). Although not demonstrated yet for AGR3, this phenomenon is likely to occur in the same manner as for AGR2 as suggested by the large number (12 out of 39) of cytosolic/nuclear proteins found associated with AGR3 in large scale protein-protein interaction experiments (54). These observations remain however to be confirmed experimentally with minimized experimental biases. Thus far the most important studies on the role of cytosolic AGR2 (cAGR2) have focused on its capacity to inhibit p53 protein tumor suppressor activity. Indeed, beyond the demonstrated genetic alterations associated with alterations of p53

functions, AGR2 could exert a non-genetic inhibition of p53 (44,55). Indeed, AGR2-p53 co-immunoprecipitation experiments performed on the cytosolic fraction extracted from A549 cells, showed that inhibition of p53 phosphorylation and activity resulted from interaction of p53 with AGR2 consecutive to a reflux of AGR2 into the cytosol in response to treatment with tunicamycin, thapsigargin, Brefeldin-A, or etoposide, a Topoisomerase II inhibitor (44,55). Moreover, preventing the association of cAGR2 with p53 by using specific AGR2 nanobodies expressed in the cytosol, resulted in the restoration of p53 activation upon stress (44,55). Cytosolic p53-dependent pro-oncogenic roles of AGR2 might also be mediated indirectly, for instance through the Dual Specificity Phosphatase 10 (DUSP-10) (56), Tumor-Susceptibility-Gene 101 (TSG-101) (53) or RuvbL2/Reptin (41), but the underlying mechanisms remain to be identified. Cytosolic roles of AGR2, and possibly AGR3 might as well (as for p53) be associated to pro-oncogenic gains-of-functions which need to be experimentally validated in relevant models (**Table 1**).

3. AGR family proteins expression in the gastrointestinal tract

3.1. Organ specificity

Analyses of the literature and databases (Human Protein Atlas – (57), (58)) showed that AGR2 mRNA and protein is expressed from the stomach to the rectum in healthy human digestive tract with the highest expression detected in the ileum and colon (59). It is mostly found in luminal mucosal cells, witnessing its secretion with mucins in the intestinal lumen (51). We confirmed that the AGR2 protein is not detected in the liver and in the two malpighian tissues of the digestive tract which are the esophagus and the anus, while it is strongly expressed in all the other parts of the digestive system including pancreas and gallbladder (**Figure 3a**). The highest expression of AGR2 is observed in both stomach, colorectal and ileum epithelia (**Figure 3a, b**). AGR1 is ubiquitously expressed in all human tissues and, in the gastrointestinal

tract it is particularly abundant in the liver and rectum (Human Protein Atlas - (26), (22)). In contrast, AGR3 is highly and almost exclusively expressed in the digestive tube, liver, respiratory system and female organs (Human Protein Atlas - (29)).

3.2. Cell specificity

AGR2 is expressed in epithelial cells from various tissues. It is essential in the production of gel-forming mucins and for the protection of the gastrointestinal tract. Cells exhibiting the highest expression of AGR2 throughout the entire gastrointestinal tract, and especially in the rectum, colon, ileum, and stomach, are mucus-secreting cells, namely goblet cells and gastric mucocytes. Undifferentiated cells, Paneth cells and enterocytes are other specialized cells with high expression of AGR2 in the digestive tube. Both Paneth and goblet cells deregulation, especially downregulation, are involved in inflammatory bowel diseases. Indeed, the loss of protective mucus layer participates to the aggression of the intestinal epithelium by pathogens and is involved in the genesis and maintenance of these diseases (60). AGR2 expression is also reported in intestinal enteroendocrine cells. In the gallbladder, biliary tract and liver, AGR2 is specifically expressed by cholangiocytes. AGR1 is ubiquitously expressed in all organs; single cell analysis show highest expression for exocrine pancreatic cells, ductal cells and pancreatic endocrine cells among all the digestive cells from a RNA single cell analysis. Regarding AGR3 mRNA expression, proximal enterocytes are the highest expressers, followed by Paneth cells, goblet cells and distal enterocytes. There is a low expression of AGR3 in enteroendocrine cells, gastric mucocytes, exocrine glandular cells. A very weak signal is found for ductal cells, cholangiocytes and pancreatic endocrine cells, while there is no expression in hepatocytes (Human Protein Atlas, Single cell - (26,57,57)).

3.3. Roles in the gastrointestinal tract

AGR2 is essential for the maintenance of homeostasis and the protection of the digestive tract, as it has a key role in mucus production, secretion, proliferation, and differentiation. AGR2 is known to interact with different mucins of the intestinal epithelial layer (MUC1, MUC2, MUC5AC, MUC5B) and is indispensable for their proper folding and secretion through its PDI-like domain (61). AGR2 could be secreted with MUC-2, but its role within the mucus layer in the intestinal lumen remains obscure (62). In the stomach, AGR2 participates to calpain production and secretion, which is indispensable for mucosa protection (28). It is also involved in mucus production in other organs such as lung (60) or pancreas (63). Mice knock-out for *Agr2* show a large intestinal hyperplasia, from the antrum to the colon, causing multiple nutrient deficiencies and a shorter lifespan compared to *Agr2*^{+/+} mice (median survival 20 weeks vs. 60 weeks) (64). In the stomach of *Agr2*^{-/-} mice, parietal and enteroendocrine cells are slightly present, and goblet cells expressing different gastric mucosal cell markers such as Trefoil factor family 2 (TFF2) and Griffonia Simplicifolia II (GSII)⁺ are overexpressed. As TFF2 and GSII marker normally disappear at birth, it is thought that AGR2 plays a key role in the maturation of gastric goblet cells. The gastric cells from *Agr2*^{-/-} mice showed increased Ki67 expression and overexpressed SOX9 responsible for the maintenance of stem cell pluripotency. All these gastric features of *Agr2*^{-/-} mice correspond to preneoplastic lesions known as spasmolytic protein expressing metaplasia (64). In the intestine, goblet cells of *Agr2*^{-/-} mice lose their normal morphology and functions, likely associated with defects in the production of mucus and MUC-2 (59). A dramatic expansion of the Paneth cell compartment and abnormal Paneth cell localization were observed in their small intestine. Mice developed a severe terminal ileo-colitis with neutrophil infiltration in the lamina propria, hyperplastic

submucosa and crypt base, Peyer's patch hyperplasia, and granulomatous inflammation. Paneth, goblet, and progenitor cells experienced ER stress (61,64,65). In conclusion, these experiments provide important information on the key roles of AGR2 in the digestive tract and especially its protective and anti-inflammatory roles.

4. AGR proteins in digestive cancers

AGR2 is overexpressed in numerous digestive solid cancers such as pancreas (66), biliary tract (67), esophagus (68), stomach (69), and colon (70,71) cancers. Analysis using the pan-cancer TCGA data from cBioportal data sets reports mainly amplifications of AGR2, but also AGR3 and AGR1 in cancers, whereas few deletions are observed (72). From a meta-analysis of 3,285 cases of different solid tumors, AGR2 is associated with the occurrence of metastasis and to a poor prognosis (73,74). Analysis of both the Human Protein Atlas and Cancer Genome Atlas (57) shows an overexpression of AGR2 at both mRNA and protein levels in most digestive neoplasia, especially gastric, colorectal, pancreatic, and biliary tract adenocarcinoma with the highest overexpression in stomach cancer. Cholangiocarcinoma (CCA) data have been grouped with those of the hepatocellular carcinoma, in which RNA and protein expression are poorly to not detected. However, a case per case analysis shows that for CCA, AGR2 is highly expressed. As AGR2 and AGR3 have similar roles in some aspects, we have compared their differential protein expression (tumors with high to moderate expression of proteins) in digestive cancers from the Human Protein Atlas (**Figure 4a, b**).¹ In intestinal and pancreatic cancers, high and moderate expression of AGR2 and AGR3 is noted in 80 to 70% of the cases, respectively. In stomach cancer, 100% of tumors have high or moderate expression of AGR2, and 80% of tumors express AGR3. Among the liver tumors of the Human Protein Atlas, one third

expresses AGR2 whereas only 1 tissue sample out of 11 express AGR3. In the liver, all tumors expressing AGR2 and AGR3 are emanating from the biliary tract (**Figure 4a, b**).

i. Esophagus adenocarcinoma

Although anomalies are very unlikely, amplification is the most recurrent anomaly in AGR2 and 3 genes in esophagus adenocarcinoma and only one mutation in AGR1 gene was observed in this cancer subtype. Regarding esophageal squamous cell carcinoma, no alterations of AGR1, 2 and 3 have been observed (20). AGR2 is overexpressed in esophagus adenocarcinoma and in Barrett metaplasia (55,75). Overexpression of AGR2 in Barrett metaplasia may be a consequence of chronic ER stress secondary to a permanent reflux of bile acid. Furthermore, immunohistochemical analysis of AGR2 has been found positive in 68 out of 69 esophagus adenocarcinoma but only focally in 15 out of 41 cases of squamous cell carcinoma, making it a marker to differentiate them in combination with other markers in the context of esophageal carcinoma of uncertain differentiation (76). Interestingly, in a whole genome array, Barrett metaplasia and esophagus adenocarcinoma tissues showed an increase of AGR2 mRNA, among other functional elements of the ER and Golgi apparatus and among secreted glycoproteins like MUC2B and MUC6, compared to healthy esophagus tissue. Among these genes, some of them are involved in the formation of a protective mucous membrane to protect the esophagus tissue against acidity. This indicates an increase of the glycosylation capacities of Barrett and esophagus adenocarcinoma compared to normal esophagus mucosa, and the potential implication of AGR2 in this increased glycosylation capacity (68).

324 ii. *Gastric adenocarcinoma*

325 Analysis of the pan-cancer TCGA data indicates that amplification of AGR2 and 3
326 genes is the most common observed alteration with very few deletions or missense
327 mutations, while nearly no missense mutations and no amplification anomaly are found
328 for AGR1 (20). Gastric adenocarcinoma cell lines, particularly SNU-520, HUG-1N,
329 MKN-45 and IM-95 are the cancer cell lines with highest expression of AGR2 in
330 cbioportal.org (20) and Depmap.org (77). A proteomic analysis comparing two gastric
331 cancer cell lines with differential potential of metastasis showed that AGR2 was the
332 most differentially overexpressed protein, supporting the capacity of AGR2 to predict
333 the risk of developing metastasis in gastric cancer (78). *In vitro* analyses carried out
334 on the gastric cancer cell line Tu-Kato III (*TP53* deleted) showed that AGR2 favors cell
335 growth and resistance to apoptosis under stress or hypoxia (79). In a murine xenograft
336 model of Tu-Kato-III treated with an AGR2 targeting miRNA, tumor growth was 8 fold
337 lower and there was much less mice with peritoneal metastases than in the control
338 group, suggesting AGR2 could contribute to gastric cancer aggressiveness and
339 metastasis (79). Expression of AGR2 has a prognostic value in gastric
340 adenocarcinoma, especially in the early stages of the disease. In a retrospective study
341 of 436 gastric cancer and 92 non-cancerous gastric samples (69) patients with low
342 expression of tumor AGR2 had a longer survival than those with high expression of
343 tumor AGR2 in TNM stages I, II and III groups, but not for those at stage IV. In this
344 study, 46,8% of gastric cancer samples had a correlation between high expression of
345 AGR2 and Cathepsin-D, a marker of tumorigenesis associated to the process of
346 metastasis (80). Increased tumor size, depth of invasion, tumor stage (based on UICC
347 TNM classification) and positive vessel invasion and lymphatic metastasis were also
348 significantly associated with high expression of AGR2 (69). However, another

immunohistochemistry analysis performed on 145 biopsies of patients with gastric adenocarcinoma didn't find a prognostic value of AGR2 (81). There was even a tendency for a poorer prognosis for patients with low expression of AGR2. This may be due to a difference in the proportion of patients with stages I-II-III vs IV. AGR1 might act as a pro-oncogenic protein in gastric cancers. Its expression was measured by immunohistochemistry and qPCR in GC tissues of patients and found to be significantly more expressed in tumor tissues than in normal tissues. Moreover, AGR1 expression was found to be correlated with tumor size and poor clinical prognosis (82).

iii. Primary liver cancers

Neither AGR2 nor AGR3 are expressed in hepatocytes and reports describing their roles in hepatocellular carcinoma (HCC) or models of HCC might be due to specificities of the samples studied. A recent study based on the analysis of transcriptomic profiles of tissue specimen of patients with HCC or intrahepatic cholangiocarcinoma (iCCA) demonstrated the existence of 4 liver cancer (LC) subtypes (LC1-LC4) showing distinct genetic variations. LC1 was a typical HCC, whereas LC2 were iCCA-like HCC, LC3 were HCC-like iCCAs and LC4 corresponded to iCCAs. Gene expression of AGR2 was absent in LC1 and appeared with the iCCA component, AGR2 being a marker of intra-hepatic CCA-like HCC associated with *TP53* mutation, ER stress and unfavorable prognosis (83). In the normal biliary tract, AGR2 is expressed in the mucus secreting tall cells that covers the large bile ducts, and its expression was found enhanced in 100% of hilar and extra-hepatic cholangiocarcinoma (eCCA) and in 50% of the iCCA (67). No data about the type of alteration are available in cbiportal.org. In addition, a cancer specific isoform of AGR2 (AGR2vH) has been identified in metastatic CCA, conferring migrating and adhesion ability to metastatic cells (84). The expression of AGR3 was analyzed on tissue sections of 74 patients diagnosed with either HCC or

iCCA by immunohistochemistry. A negative or very weak heterogeneously expression of AGR3 was observed in HCC patients, whereas AGR3 expression was moderate or strong in iCCA tissues. Thus, AGR3 was proposed to be used as a complementary marker in liver cancers to distinguish between HCC and iCCA for diagnosis (85). Fibrolamellar carcinoma (FLHCC), is a rare liver cancer that can arise in young patients without any previous liver disease. AGR2 was found often overexpressed in FLHCC (86). At last, 75% of primary FLHCC and 75% of metastatic FLHCC were positive for AGR2 (87). At last, AGR1 protein expression was evaluated in 106 human HCC tissues by immunohistochemistry and was significantly higher compared to corresponding adjacent normal tissues. In addition, its expression was correlated with HCC metastasis and AGR1 was found to promote metastasis through the induction of the epithelial–mesenchymal transition process (88), however the precise underlying molecular mechanisms remain unclear.

iv. Pancreatic cancer

Few amplifications of AGR2 and AGR3 genes (<1%) have been demonstrated in pancreatic adenocarcinoma as well as very few missense mutations (<1%) have been observed for AGR2 (20). So far, no alterations have been demonstrated for the AGR1 gene. Prolonged chronic ER stress induces the overexpression of AGR2 in normal Human Pancreatic Duct Epithelial (HPDE) cell line (66), and may be involved in development of pancreatic adenocarcinoma. In *LSL-Kras^{G12D};Pdx1-Cre* (KC) mouse, the induction of AGR2 expression precedes the formation of PanIN, whereas PanIN do not develop in *Agr2^{-/-}* mice, suggesting the involvement of AGR2 in the development of pancreatic neoplasms (66). In human, AGR2 is highly expressed in PanIN compared to normal pancreatic tissue, and its expression is higher in PanIN than in invasive pancreatic adenocarcinoma (45,89–91). According to a retrospective

analysis of 57 biopsy samples, AGR2 mRNA expression is 14 times higher in pancreatic adenocarcinoma than normal pancreatic or pancreatitis tissues, with similar tendency for AGR2 protein expression (92). These data thus suggest that AGR2 overexpression caused by ER stress is an early event in pancreatic cancer genesis and may be important for the progression from normal pancreas tissue to invasive pancreatic cancer. The prognostic value of AGR2 in pancreatic cancer is debated, with discordant results (89,90). A high expression was found to correlate with poor prognosis in one study (89), whereas in another study, a low expression related to poor prognosis only using univariate analysis, but could not be confirmed using multivariate analysis (91). However, a high expression of AGR2 may be correlated to poor differentiation of pancreatic ductal adenocarcinoma, and therefore to pancreatic cancer aggressiveness (91). In pancreatic cancer cell line MPanc-96, AGR2 is involved in cell proliferation and invasion, resistance to gemcitabine and to apoptosis. In an *in vivo* murine orthotopic pancreas tumor model, the silencing of AGR2 resulted in a 4 fold reduction of tumor volume compared to control, tumor sensitization to gemcitabine, and fewer metastases in the group treated with gemcitabine and silenced for AGR2 (92). Interestingly, cell proliferation, migration and chemoresistance in different cultured pancreatic cells may be favored by the soluble form of AGR2 (93). All these data suggest AGR2 is involved in pancreas cancer proliferation and chemoresistance.

v. Colorectal cancer

AGR2 is one of the most overexpressed proteins in human colorectal cancer (CRC), especially in mucinous forms (94). In CRC, the expression and prognostic role of AGR2 remains to be fully understood. Indeed, Tian et al. (74), performed a comprehensive meta-analysis using data of 2 colorectal cancer studies (70,95) selected based on their high methodological quality (Newcastle-Ottawa Scale (NOS) scores between 8 and 9).

424 For the first study, blood samples of 54 stages I-IV CRC patients were collected and
425 AGR2 expression was analyzed by qPCR. High and low AGR2 expressers were
426 defined based on the mean relative AGR2 mRNA expression (high expressers > mean
427 expression level and low expresser $\leq$ mean expression level). Independently, this
428 study found that AGR2 expression was correlated with stage III and high-grade tumors.
429 In addition, progression-free survival (PFS) of patients with high AGR2 expression was
430 decreased (Hazard Ratio: 2.32; $p = 0.037$) (95). For the second study, AGR2
431 expression was measured by immunohistochemistry in 432 CRC cases and a scoring
432 system was used to make the classification based on the staining intensity: negative
433 (0), weakly (1), moderately (2) and strongly positive (3). Low and high expressers were
434 defined according to their AGR2 expression as either low (score 0 or 1) or high (score
435 2 or 3). Low expression of AGR2 in CRC tissue samples was significantly associated
436 with high tumor grade, left primary tumor localization as well as shorter patient survival
437 times. 78% of patients had a lower AGR2 tumor expression compared to normal
438 colonic mucosa (70). Interestingly, when combined together, pooled hazard ratio
439 between the two studies, AGR2 overexpression could not predict poor outcomes in
440 colorectal cancer: Hazard Ratio 0.80, 95% Confidence Interval: 0.41–1.53 (74). CRC
441 cell proliferation and metastatic process may be favored by the secreted form of AGR2
442 (71). For example, mice injected with CRC cell line HT-29 *Agr2*^{-/-} developed
443 significantly more metastases after treatment with eAGR2 than control. In different
444 CRC cell lines, it has also been shown that eAGR2 can inhibit the canonical Wnt/b-
445 catenin pathway (involved in proliferation) and stimulate the non-canonical Wnt/ b-
446 catenin CAMKII-JNK pathway by promoting Wnt11 overexpression, which is a
447 regulator of cell polarity and motility, stimulating the metastatic process (71). In induced
448 CRC stem cells, the canonic Wnt/b-catenin pathway stimulates AGR2 expression

through b-catenin binding to the *AGR2* promoter (96). *AGR2* could also be involved in maintaining stem cells stemness potential since its silencing using shRNA induced the loss of their stem cell characteristics (96). The implication of *AGR2* in maintaining cancer stem cell potential has also been reported in head and neck squamous-cell carcinoma cells (97). *AGR2* might act as a tumor suppressor in CRC tumors however, future experimental verifications in bigger sample size studies are required. The functional role of *AGR3* in CRC was investigated in 8 CRC tissue samples by immunohistochemistry and 24 CRC tumor tissues by western blot. *AGR3* was significantly more expressed in CRC tumors tissues than in the adjacent normal tissues for the 8 patients. In addition, western blot results highlighted higher expression of *AGR3* in cancer tissue in 13 out of 24 tumors compared with matched adjacent normal tissues. High expression of *AGR3* was also significantly associated with poorer overall survival and *AGR3* expression was identified as an independent prognostic factor for poor survival (98). Although the results regarding the role of *AGR3* in CRC are quantitatively insufficient and need further experimental validation, they suggest an overexpression of *AGR3* in CRC which is associated with poor prognosis.

5. AGR proteins in IBD and colitis associated cancer

IBD are characterized by deficiencies in mucus layer that is part of the intestinal epithelial barrier leading to incomplete protection against microbial aggression. Consequently, excessive immune responses are triggered which results in chronic inflammation. *AGR2* has been shown to be crucial for the protection of the mucus barrier in several studies. Whole genome sequencing performed in two siblings presenting infantile IBD, revealed a new *AGR2* variant which results in *AGR2* loss of function. *AGR2* deficiency led to increased ER stress and a decrease of MUC2

processing which impairs the proper formation of the mucus barrier and thus induces intestinal inflammation (99). AGR2 expression was measured by qPCR in small and large bowel biopsies of 56 Crohn's disease (CD) and 57 ulcerative colitis (UC) patients and shown to be significantly lower in both disease groups compared to 25 healthy controls. Unfortunately, this study does not specify if the patients were undergoing treatment at the time of the study (100). In contrast, AGR2 gene expression was significantly upregulated in rectal biopsies of active patients with UC compared to both the remission and control groups. However, AGR2 gene expression might be modulated with the use of anti-inflammatory drugs and AGR2 higher expression in active patients could be a consequence of therapeutic improvement (101). In addition, AGR2 monomers homodimers ratio homeostasis was shown to be altered in CD. This imbalance results in AGR2 secretion leading to monocytes recruitment and inflammation (34). Patients with IBD are 2 to 3 times more likely to develop CRC cancer than the general population. They are also at increased risk for CRC death (102,103). Mice genetically deficient in Muc2 developed tumors in the small and large intestine including the rectum (104). AGR2 is essential for the production of intestinal MUC2 (61) indicating a critical role of AGR2 expression for protection to colitis associated cancer. To our knowledge, the contribution of AGR1 and AGR3 in IBD has never been demonstrated.

6. Theranostic perspectives

AGR2 has a physiological role in ER and interacts with a number of cellular partners depending on the cell type considered. Fourteen AGR2-interacting proteins overlapped in three independent studies (50,105,106). These proteins are involved predominantly in processes into the ER which contribute to maintaining of intracellular metabolic

homeostasis and/or to responses to ER stress (UPR, protein folding, changes in cellular metabolism, and redox state). Given the importance of the AGR2 interaction network in maintaining cellular homeostasis, the specific targeting of ER-resident AGR2 by potential therapeutic tools is difficult to envisage, as demonstrated by the deleterious effects observed in mice invalidated for this gene. On the contrary, secreted or cytosolic forms of AGR2 that play pro-oncogenic and pro-inflammatory functions appear to be more relevant to consider and open new avenues from a therapeutic point of view.

i Development of monoclonal antibodies and immunotherapies

Monoclonal antibodies (mAbs) directed against AGR2 have already been experimented in animal models with promising results. The main therapeutic objective is to specifically target the extracellular AGR2, without impacting the ER-resident protein. Promising results have been reported in experimental models. 18A4Hu, an anti-AGR2 mAb, reduced tumor volume and tumor weight as much as Bevacizumab, without significant side effects, in a mammary breast cancer xenograft model (107). However, the combination with Bevacizumab was not studied, which would have been interesting as we already know that AGR2 and VEGF-A act in synergy (71,108–110). The 18A4Hu mAb is also effective on lung cancer (A549 and H460) and melanoma (B16F10) cell lines by inhibiting tumor cell growth and migration, disrupting cell morphology and junctions, and increasing p53 and p21 expression (111). A murine xenograft model of A549 and H460 cells treated with 18A4Hu showed a 3 times fold reduction of tumor volume and tumor weight compared to untreated mice. Tumor analysis of treated mice showed an increase in apoptotic markers and p53 expression, and a reduction of tumor blood vessels compared to untreated mice, suggesting that 18A4Hu has antiangiogenic properties. Similarly, mice with lung metastasis of

524 melanoma cell line B16F10, treated with 18A4Hu exhibits 10 times less metastasis in
525 lungs, and their survival was significantly improved (24 vs. 20 days). Again, it was well
526 tolerated, and there was neither reduction of intestinal mucus nor occurrence of
527 intestinal inflammation. P1G4 and P3A5 are two other anti-AGR2 humanized mAbs.
528 P1G4 is not efficient in monotherapy on a xenograft model of pancreatic
529 adenocarcinoma (112). However, its association with gemcitabine decreased
530 significantly tumor volume and increased survival compared to gemcitabine alone.
531 P3A5 is not efficient alone or combined with gemcitabine to reduce tumor volume but
532 it increased survival in combination with gemcitabine. A combination of mouse mAbs
533 directed against AGR2 and C4.4A in a murine xenograft model of pancreatic cancer
534 resulted in a 50% objective response rate, a 50% reduction of the occurrence of
535 metastasis, and an increase in OS. Their efficacy was similar or better than
536 gemcitabine alone or combined with the two mAbs (93). These preliminary results are
537 promising as they prove that mAbs anti-AGR2 have *in vivo* effects in diverse cancers,
538 and should be investigated further. Therapeutic progresses in digestive oncology
539 remains challenging since, with the exception of immunotherapy in microsatellite
540 instability (MSI) or deficient Mismatch Repair (MMR) CRC and oeso-gastric cancers,
541 most of the targeted therapies currently used have been developed more than ten
542 years ago and no other important progresses in terms of survival have been made. In
543 addition to the development of anti-AGR2 therapeutic antibodies, dendritic cell (DC)
544 vaccines with AGR2 as a potent antigen could improve cancer immunotherapy. DCs
545 were transduced with a recombinant adenovirus encoding AGR2 and these
546 engineered DCs both increased the number of T cells secreting IFN γ and induced the
547 lysis of AGR2-expressing colorectal cancer cell lines through the activation of potent
548 AGR2-specific cytotoxic T cells (113). Although limitations persist and impede the

application of DCs vaccines immunotherapy, these in vitro experiments highlight the potential of AGR2 to trigger efficient immune responses against some cancers. Further research in animals is required to validate the therapeutic relevance of these findings.

ii AGR2 and AGR3 as biomarkers?

The identification of biomarkers using non-invasive sampling in digestive cancer is crucial to improve efficacy of diagnostic, predict prognosis and treatment outcome (or clinical management of patients). AGR2 gene expression has been assessed in whole blood of CRC patients before therapy by qPCR. AGR2 mRNA expression was significantly increased in the blood of patients with CRC compared to controls. In addition, the level of AGR2 gene expression was correlated with some invasive and high-grade tumor prognostic factors (95). Thus, high AGR2 gene expression in the blood of CRC patients might be a sign of poor patient outcome. However, this would need to be confirmed with a longitudinal study, while its relevance in therapeutic strategy would remain to be defined. AGR2 protein level has been measured in the plasma of pancreatic cancer patients before treatment using ELISA and is significantly higher compared to healthy controls (114). A diagnostic score comprising CA19.9, AGR2 and Regenerating islet-derived 1 beta (REG1B) serum levels was shown to be more efficient than CA19.9 dosage alone in differentiating pancreatic ductal adenocarcinoma (PDAC) from benign pancreatic tumor or healthy controls, making it a potential interesting score to avoid systematic biopsies in pancreatic tumor, of which are often difficult to perform with a diagnostic performance not optimal, which often necessitates to perform additional biopsies. However, this score has not been confirmed by a prospective study. In addition, AGR2 was significantly more elevated in the serum of patients with PDAC than in other tumors, while AGR2 alone was the

best marker for pancreas cancer diagnostic after CA19.9 (114). Although these results need to be validated in larger cohorts of samples, AGR2, in combination with other biomarkers, could be used for better patient management through early disease detection. To improve cancer diagnosis (/for early cancer detection), several research teams have developed high sensitive assays for rapid and ultrasensitive detection of AGR2 protein (115,116). Nevertheless, the application of these highly sensitive assays to plasma samples needs to be evaluated for non-invasive early detection of cancer. However, as AGR2 is not specific to one of the digestive organs, it cannot be used alone for early cancer detection as a serum biomarker, and it needs to be integrated in diagnostic scores. Overall, very few studies have investigated AGR2 protein level in the serum of digestive cancer patients. AGR2 might be a good candidate to predict clinical outcome such as metastatic potential or relapse but it needs to be demonstrated with longitudinal studies in cohort of patients. There is no data regarding AGR2 efficacy as treatment outcome predictor in digestive cancers. Moreover, the presence of AGR3 in the serum of patients with digestive cancer has never been explored. There are also limited data on AGR3 expression in healthy and digestive cancer tissues. Only one study investigated the potential role of AGR3 as a diagnostic biomarker to distinguish between iCCA and HCC in digestive cancer (as previously described in the primary liver cancers part of this review). AGR3 expression was measured in tissues of iCCA and HCC patients by immunohistochemistry and AGR3 was found to be expressed only by intrahepatic bile duct cholangiocytes but not hepatocytes (91). Thus, AGR3 could be used as a new biomarker for differential diagnosis between these two cancers.

7 Conclusion

The AGR2 protein has gained significant interest in oncology during the last years. This holds to its multiple pro-oncogenic properties and to the particularity that its pathological roles mostly depend on its cellular or extracellular localization. AGR2 is highly expressed in the digestive system. It is overexpressed at early stages in most cancer lesions and adenocarcinoma of the digestive tube and is often associated to a poor prognosis. Less studied than in cancer, AGR2 deregulation may also play an important role in inflammation, particularly in inflammatory bowel diseases (100,117). However, the role of extracellular AGR2 in the tumor microenvironment, its influence on the immune infiltrate, its autocrine and paracrine roles remain to be more clearly characterized. The development of anti-AGR2 therapies has already shown interesting results in animals and plasma/serum AGR2 could serve as prognostic/diagnostic biomarker. In the future, their development may have prospects in diverse pathologies like cancer and inflammatory diseases and may enter into the therapeutic arsenal of the treatment of these patients.

Acknowledgements and Funding: We thank the Immuno-histopathology platform H2P2 for their expertise and work (<https://histopathologie.univ-rennes1.fr/>) and Raphael Pineau for his technical help. This work was funded by grants from INCa (PRT-K20-136) to CC and JE, INCa (PLBIO) to EC, La Ligue contre le cancer to EOD and AL and from La Ligue Contre le Cancer Gironde and from the Site de recherche intégrée sur le cancer de Bordeaux (SIRIC Brio) to FD. LAE acknowledges funding from the Swedish research council (grant no 2019-3684) and the Swedish Cancer foundation (grant no 21-1447 Pj). EB was funded by an “année-recherche” grant from the ministry of health.

Competing interests: EOD, XT and EC are founders of Thabor Therapeutics

Patient consent for publication: According to the French regulation, patients did not oppose to the analysis, and the study was approved by the Rennes Ethics Committee (Avis n° 21.122).

ORCID iDs:	EB	0000-0003-1831-6145
	CP	0000-0003-4182-6442
	FDM	0000-0001-8201-7997
	JE	0000-0002-8289-7741
	CC	0000-0002-5692-9586
	RH --	0000-0002-6139-2664
	AM --	0000-0002-0081-0498
	FD	0000-0002-4600-7633
	XT	0000-0001-8123-9213

639 LAE 0000-0001-5654-3109
640 EC 0000-0001-5855-4522
641 AL 0000-0003-2118-717X
642 EOD 0000-0002-0057-7593

643

644 **Author contribution statement**

645 EB, EC, EOD and CP wrote the first draft of the manuscript, made the figures and
646 finalized the document. LAE carried out the structural analyses presented in Figure 1.
647 CP, FDM, JE, RH, AM, CC, FD, XT, LAE, AL critically read the manuscript and worked
648 on it.

649

References

1. Chevet E, Fessart D, Delom F, Mulot A, Vojtesek B, Hrstka R, et al. Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development. *Oncogene*. 2013 May;32(20):2499–509.
2. Määttänen P, Gehring K, Bergeron JJM, Thomas DY. Protein quality control in the ER: the recognition of misfolded proteins. *Semin Cell Dev Biol*. 2010 Jul;21(5):500–11.
3. Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE. The folding catalyst protein disulfide isomerase is constructed of active and inactive thioredoxin modules. *Curr Biol CB*. 1997 Apr 1;7(4):239–45.
4. Matsusaki M, Kanemura S, Kinoshita M, Lee YH, Inaba K, Okumura M. The Protein Disulfide Isomerase Family: from proteostasis to pathogenesis. *Biochim Biophys Acta Gen Subj*. 2020 Feb;1864(2):129338.
5. Lemaire SD, Miginiac-Maslow M. The thioredoxin superfamily in *Chlamydomonas reinhardtii*. *Photosynth Res*. 2004;82(3):203–20.
6. Matsuo Y. Introducing Thioredoxin-Related Transmembrane Proteins: Emerging Roles of Human TMX and Clinical Implications. *Antioxid Redox Signal*. 2022 May;36(13–15):984–1000.
7. Kozlov G, Määttänen P, Thomas DY, Gehring K. A structural overview of the PDI family of proteins. *FEBS J*. 2010;277(19):3924–36.

- 670 8. Ivanova AS, Tereshina MB, Ermakova GV, Belousov VV, Zارايسكى AG. Agr
671 genes, missing in amniotes, are involved in the body appendages regeneration
672 in frog tadpoles. *Sci Rep*. 2013;3:1279.
- 673 9. Fessart D, Mahouche I, Brouste V, Velasco V, Soubeyran I, Soubeyran P, et al.
674 Anterior Gradient-2 (AGR2) overexpression in colon cancer: a potential
675 prognostic biomarker [Internet]. *bioRxiv*; 2021 [cited 2022 Jun 5]. p.
676 2021.09.07.459258. Available from:
677 <https://www.biorxiv.org/content/10.1101/2021.09.07.459258v1>
- 678 10. Higa A, Mulot A, Delom F, Bouche-careilh M, Nguyễn DT, Boismenu D, et al.
679 Role of pro-oncogenic protein disulfide isomerase (PDI) family member anterior
680 gradient 2 (AGR2) in the control of endoplasmic reticulum homeostasis. *J Biol*
681 *Chem*. 2011 Dec 30;286(52):44855–68.
- 682 11. Hatahet F, Ruddock LW. Protein Disulfide Isomerase: A Critical Evaluation of Its
683 Function in Disulfide Bond Formation. *Antioxid Redox Signal*. 2009 May
684 28;11(11):2807–50.
- 685 12. Ni M, Lee AS. ER chaperones in mammalian development and human diseases.
686 *FEBS Lett*. 2007 Jul 31;581(19):3641–51.
- 687 13. Fessart D, Villamor I, Chevet E, Delom F, Robert J. Integrative analysis of
688 genomic and transcriptomic alterations of AGR2 and AGR3 in cancer. *Open*
689 *Biol*. 12(7):220068.
- 690 14. Aberger F, Weidinger G, Grunz H, Richter K. Anterior specification of embryonic
691 ectoderm: the role of the *Xenopus* cement gland-specific gene XAG-2. *Mech*
692 *Dev*. 1998 Mar;72(1–2):115–30.

15. Park SW, Zhen G, Verhaeghe C, Nakagami Y, Nguyenvu LT, Barczak AJ, et al.
The protein disulfide isomerase AGR2 is essential for production of intestinal
mucus. *Proc Natl Acad Sci U S A*. 2009 Apr 28;106(17):6950–5.
16. Shih LJ, Lu YF, Chen YH, Lin CC, Chen JA, Hwang SPL. Characterization of the
agr2 gene, a homologue of *X. laevis* anterior gradient 2, from the zebrafish,
Danio rerio. *Gene Expr Patterns GEP*. 2007 Feb;7(4):452–60.
17. Bradley L, Wainstock D, Sive H. Positive and negative signals modulate
formation of the *Xenopus* cement gland. *Dev Camb Engl*. 1996
Sep;122(9):2739–50.
18. Jach D, Cheng Y, Prica F, Dumartin L, Crnogorac-Jurcevic T. From
development to cancer - an ever-increasing role of AGR2. *Am J Cancer Res*.
2021;11(11):5249–62.
19. Petek E, Windpassinger C, Egger H, Kroisel PM, Wagner K. Localization¹ of the
human anterior gradient-2 gene (AGR2) to chromosome band 7p21.3 by
radiation hybrid mapping and fluorescencein situ hybridisation. *Cytogenet
Genome Res*. 2000;89(3–4):141–2.
20. cBioPortal for Cancer Genomics [Internet]. [cited 2021 Jun 14]. Available from:
<https://www.cbioportal.org/>
21. Knoblach B, Keller BO, Groenendyk J, Aldred S, Zheng J, Lemire BD, et al.
ERp19 and ERp46, New Members of the Thioredoxin Family of Endoplasmic
Reticulum Proteins*. *Mol Cell Proteomics*. 2003 Oct 1;2(10):1104–19.

- 714 22. Williamson RA, Howard MJ, Jääntti HP, Rautio SM, Kellokumpu S, Alanen HI, et
715 al. Functional Characterization of ERp18, a New Endoplasmic Reticulum-located
716 Thioredoxin Superfamily Member*. J Biol Chem. 2003 Aug 1;278(31):28912–20.
- 717 23. Ivanova AS, Shandarin IN, Ermakova GV, Minin AA, Tereshina MB, Zaraisky
718 AG. The secreted factor Ag1 missing in higher vertebrates regulates fins
719 regeneration in *Danio rerio*. Sci Rep. 2015 Jan 29;5:8123.
- 720 24. Alanen HI, Williamson RA, Howard MJ, Lappi AK, Jääntti HP, Rautio SM, et al.
721 Functional characterization of ERp18, a new endoplasmic reticulum-located
722 thioredoxin superfamily member. J Biol Chem. 2003 Aug 1;278(31):28912–20.
- 723 25. Jeong W, Lee DY, Park S, Rhee SG. ERp16, an endoplasmic reticulum-resident
724 thiol-disulfide oxidoreductase: biochemical properties and role in apoptosis
725 induced by endoplasmic reticulum stress. J Biol Chem. 2008 Sep
726 12;283(37):25557–66.
- 727 26. TXNDC12 protein expression summary - The Human Protein Atlas [Internet].
728 [cited 2022 Jul 25]. Available from:
729 <https://www.proteinatlas.org/ENSG00000117862-TXNDC12>
- 730 27. Raykhel I, Alanen H, Salo K, Jurvansuu J, Nguyen VD, Latva-Ranta M, et al. A
731 molecular specificity code for the three mammalian KDEL receptors. J Cell Biol.
732 2007 Dec 17;179(6):1193–204.
- 733 28. Obacz J, Takacova M, Brychtova V, Dobes P, Pastorekova S, Vojtesek B, et al.
734 The role of AGR2 and AGR3 in cancer: Similar but not identical. Eur J Cell Biol.
735 2015 Mar 1;94(3):139–47.

29. AGR3 protein expression summary - The Human Protein Atlas [Internet]. [cited 2022 Jul 25]. Available from: <https://www.proteinatlas.org/ENSG00000173467-AGR3>
30. Adam PJ, Boyd R, Tyson KL, Fletcher GC, Stamps A, Hudson L, et al. Comprehensive Proteomic Analysis of Breast Cancer Cell Membranes Reveals Unique Proteins with Potential Roles in Clinical Cancer*. *J Biol Chem*. 2003 Feb 21;278(8):6482–9.
31. Fletcher GC, Patel S, Tyson K, Adam PJ, Schenker M, Loader JA, et al. hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan. *Br J Cancer*. 2003 Feb;88(4):579–85.
32. Bonser LR, Schroeder BW, Ostrin LA, Baumlin N, Olson JL, Salathe M, et al. The Endoplasmic Reticulum Resident Protein AGR3. Required for Regulation of Ciliary Beat Frequency in the Airway. *Am J Respir Cell Mol Biol*. 2015 Mar 9;53(4):536–43.
33. Nguyen VD, Biterova E, Salin M, Wierenga RK, Ruddock LW. Crystal structure of human anterior gradient protein 3. *Acta Crystallogr Sect F Struct Biol Commun*. 2018 Jul 1;74(7):425–30.
34. Maurel M, Obacz J, Avril T, Ding YP, Papadodima O, Treton X, et al. Control of anterior GRadient 2 (AGR2) dimerization links endoplasmic reticulum proteostasis to inflammation. *EMBO Mol Med*. 2019 Jun 1;11(6):e10120.
35. Patel P, Clarke C, Barraclough DL, Jowitt TA, Rudland PS, Barraclough R, et al. Metastasis-Promoting Anterior Gradient 2 Protein Has a Dimeric Thioredoxin

- 759 Fold Structure and a Role in Cell Adhesion. *J Mol Biol.* 2013 Mar 11;425(5):929–
760 43.
- 761 36. Delom F, Mohtar MA, Hupp T, Fessart D. The anterior gradient-2 interactome.
762 *Am J Physiol Cell Physiol.* 2020 Jan 1;318(1):C40–7.
- 763 37. Clarke DJ, Murray E, Faktor J, Mohtar A, Vojtesek B, MacKay CL, et al. Mass
764 spectrometry analysis of the oxidation states of the pro-oncogenic protein
765 anterior gradient-2 reveals covalent dimerization via an intermolecular disulphide
766 bond. *Biochim Biophys Acta.* 2016 May;1864(5):551–61.
- 767 38. Černocká H, Vonka P, Kasalová V, Sommerova L, Vandova V, Hrstka R, et al.
768 AGR2-AGR3 hetero-oligomeric complexes: Identification and characterization.
769 *Bioelectrochemistry Amst Neth.* 2021 Aug;140:107808.
- 770 39. Fomenko DE, Gladyshev VN. CxxS: Fold-independent redox motif revealed by
771 genome-wide searches for thiol/disulfide oxidoreductase function. *Protein Sci*
772 *Publ Protein Soc.* 2002 Oct;11(10):2285–96.
- 773 40. Gray TA, MacLaine NJ, Michie CO, Bouchalova P, Murray E, Howie J, et al.
774 Anterior Gradient-3: A novel biomarker for ovarian cancer that mediates cisplatin
775 resistance in xenograft models. *J Immunol Methods.* 2012 Apr 30;378(1):20–32.
- 776 41. Maslon MM, Hrstka R, Vojtesek B, Hupp TR. A Divergent Substrate-Binding
777 Loop within the Pro-oncogenic Protein Anterior Gradient-2 Forms a Docking Site
778 for Reptin. *J Mol Biol.* 2010 Dec 3;404(3):418–38.
- 779 42. Oka OBV, Pierre AS, Pringle MA, Tungum W, Cao Z, Fleming B, et al.
780 Activation of the UPR sensor ATF6 α is regulated by its redox-dependent

781 dimerization and ER retention by ERp18. *Proc Natl Acad Sci U S A*. 2022 Mar
782 22;119(12):e2122657119.

783 43. Oka OB, van Lith M, Rudolf J, Tungkum W, Pringle MA, Bulleid NJ. ERp18
784 regulates activation of ATF6 α during unfolded protein response. *EMBO J*. 2019
785 Aug 1;38(15):e100990.

786 44. Sicari D, Centonze FG, Pineau R, Le Reste PJ, Negroni L, Chat S, et al. Reflux
787 of Endoplasmic Reticulum proteins to the cytosol inactivates tumor suppressors.
788 *EMBO Rep*. 2021 May 5;22(5):e51412.

789 45. Dumartin L, Whiteman HJ, Weeks ME, Hariharan D, Dmitrovic B, Iacobuzio-
790 Donahue CA, et al. AGR2 Is a Novel Surface Antigen That Promotes the
791 Dissemination of Pancreatic Cancer Cells through Regulation of Cathepsins B
792 and D. *Cancer Res*. 2011 Nov 15;71(22):7091–102.

793 46. CALU calumenin [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2021
794 Jan 22]. Available from: <https://www.ncbi.nlm.nih.gov/gene/813>

795 47. RCN1 reticulocalbin 1 [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited
796 2021 Jan 22]. Available from: <https://www.ncbi.nlm.nih.gov/gene/5954>

797 48. Honoré B, Vorum H. The CREC family, a novel family of multiple EF-hand, low-
798 affinity Ca²⁺-binding proteins localised to the secretory pathway of mammalian
799 cells. *FEBS Lett*. 2000;466(1):11–8.

800 49. Mohtar MA, Hernychova L, O'Neill JR, Lawrence ML, Murray E, Vojtesek B, et
801 al. The Sequence-specific Peptide-binding Activity of the Protein Sulfide

Isomerase AGR2 Directs Its Stable Binding to the Oncogenic Receptor EpCAM.
Mol Cell Proteomics MCP. 2018 Apr;17(4):737–63.

50. Bouchalova P, Sommerova L, Potesil D, Martisova A, Lapcik P, Koci V, et al.
Characterization of the AGR2 Interactome Uncovers New Players of Protein
Disulfide Isomerase Network in Cancer Cells. Mol Cell Proteomics MCP. 2022
Feb;21(2):100188.

51. Bergström JH, Berg KA, Rodríguez-Piñeiro AM, Stecher B, Johansson MEV,
Hansson GC. AGR2, an Endoplasmic Reticulum Protein, Is Secreted into the
Gastrointestinal Mucus. PLOS ONE. 2014 Aug 11;9(8):e104186.

52. Moidu NA, A Rahman NS, Syafruddin SE, Low TY, Mohtar MA. Secretion of pro-
oncogenic AGR2 protein in cancer. Heliyon. 2020 Sep 1;6(9):e05000.

53. Gray TA, Alsamman K, Murray E, Sims AH, Hupp TR. Engineering a synthetic
cell panel to identify signalling components reprogrammed by the cell growth
regulator anterior gradient-2. Mol Biosyst. 2014 Jun;10(6):1409–25.

54. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID:
a general repository for interaction datasets. Nucleic Acids Res. 2006 Jan
1;34(Database issue):D535-539.

55. Pohler E, Craig AL, Cotton J, Lawrie L, Dillon JF, Ross P, et al. The Barrett's
Antigen Anterior Gradient-2 Silences the p53 Transcriptional Response to DNA
Damage. Mol Cell Proteomics. 2004 Jun 1;3(6):534–47.

- 822 56. Hrstka R, Bouchalova P, Michalova E, Matoulkova E, Muller P, Coates PJ, et al.
823 AGR2 oncoprotein inhibits p38 MAPK and p53 activation through a DUSP10-
824 mediated regulatory pathway. *Mol Oncol*. 2016 May;10(5):652–62.
- 825 57. AGR2 protein expression summary - The Human Protein Atlas [Internet]. [cited
826 2022 Jul 25]. Available from: [https://www.proteinatlas.org/ENSG00000106541-](https://www.proteinatlas.org/ENSG00000106541-AGR2)
827 AGR2
- 828 58. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et
829 al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015 Jan
830 23;347(6220):1260419.
- 831 59. Zhao F, Edwards R, Dizon D, Afrasiabi K, Mastroianni JR, Geyfman M, et al.
832 Disruption of Paneth and goblet cell homeostasis and increased endoplasmic
833 reticulum stress in *Agr2*^{-/-} mice. *Dev Biol*. 2010 Feb 15;338(2):270–9.
- 834 60. Schroeder BW, Verhaeghe C, Park SW, Nguyenvu LT, Huang X, Zhen G, et al.
835 AGR2 Is Induced in Asthma and Promotes Allergen-Induced Mucin
836 Overproduction. *Am J Respir Cell Mol Biol*. 2012 Aug 1;47(2):178–85.
- 837 61. Park SW, Zhen G, Verhaeghe C, Nakagami Y, Nguyenvu LT, Barczak AJ, et al.
838 The protein disulfide isomerase AGR2 is essential for production of intestinal
839 mucus. *Proc Natl Acad Sci*. 2009 Apr 28;106(17):6950–5.
- 840 62. Pelaseyed T, Bergström JH, Gustafsson JK, Ermund A, Birchenough GMH,
841 Schütte A, et al. The mucus and mucins of the goblet cells and enterocytes
842 provide the first defense line of the gastrointestinal tract and interact with the
843 immune system. *Immunol Rev*. 2014 Jul;260(1):8–20.

- 844 63. Norris AM, Gore A, Balboni A, Young A, Longnecker DS, Korc M. AGR2 is a
845 SMAD4-suppressible gene that modulates MUC1 levels and promotes the
846 initiation and progression of pancreatic intraepithelial neoplasia. *Oncogene*.
847 2013 Aug 15;32(33):3867–76.
- 848 64. Gupta A, Wodziak D, Tun M, Bouley DM, Lowe AW. Loss of Anterior Gradient 2
849 (Agr2) Expression Results in Hyperplasia and Defective Lineage Maturation in
850 the Murine Stomach. *J Biol Chem*. 2013 Aug 2;288(6):4321–33.
- 851 65. Li J, Hu J, Luo Z, Zhou C, Huang L, Zhang H, et al. AGR2 is controlled by
852 DNMT3a-centered signaling module and mediates tumor resistance to 5-Aza in
853 colorectal cancer. *Exp Cell Res*. 2019 Dec 1;385(1):111644.
- 854 66. Dumartin L, Alrawashdeh W, Trabulo SM, Radon TP, Steiger K, Feakins RM, et
855 al. ER stress protein AGR2 precedes and is involved in the regulation of
856 pancreatic cancer initiation. *Oncogene*. 2017 01;36(22):3094–103.
- 857 67. Lepreux S, Bioulac-Sage P, Chevet E. Differential expression of the anterior
858 gradient protein-2 is a conserved feature during morphogenesis and
859 carcinogenesis of the biliary tree. *Liver Int*. 2011;31(3):322–8.
- 860 68. Nancarrow DJ, Clouston AD, Smithers BM, Gotley DC, Drew PA, Watson DI, et
861 al. Whole genome expression array profiling highlights differences in mucosal
862 defense genes in Barrett's esophagus and esophageal adenocarcinoma. *PLoS*
863 *One*. 2011;6(7):e22513.
- 864 69. ZHANG J, JIN Y, XU S, ZHENG J, ZHANG Q, WANG Y, et al. AGR2 is
865 associated with gastric cancer progression and poor survival. *Oncol Lett*. 2016
866 Mar;11(3):2075–83.

70. Riener MO, Thiesler T, Hellerbrand C, Amann T, Cathomas G, Fritzsche FR, et al. Loss of Anterior Gradient-2 expression is an independent prognostic factor in colorectal carcinomas. *Eur J Cancer*. 2014 Jul 1;50(10):1722–30.
71. Tian S, Hu J, Tao K, Wang J, Chu Y, Li J, et al. Secreted AGR2 promotes invasion of colorectal cancer cells via Wnt11-mediated non-canonical Wnt signaling. *Exp Cell Res*. 2018 Mar 15;364(2):198–207.
72. cBioPortal for Cancer Genomics: AGR2, AGR3 and 1 other gene in Pan-cancer analysis of whole genomes (ICGC/TCGA, Nature 2020) [Internet]. [cited 2022 Jul 16]. Available from: https://www.cbioportal.org/results?cancer_study_list=pancan_pcaWG_2020&tab_index=tab_visualize&case_set_id=pancan_pcaWG_2020_all&Action=Submit&gene_list=AGR2%25AAGR3%25ATXNDC12
73. Barraclough DL, Platt-Higgins A, de Silva Rudland S, Barraclough R, Winstanley J, West CR, et al. The metastasis-associated anterior gradient 2 protein is correlated with poor survival of breast cancer patients. *Am J Pathol*. 2009 Nov;175(5):1848–57.
74. Tian S bo, Tao K xiong, Hu J, Liu Z bo, Ding X liang, Chu Y nan, et al. The prognostic value of AGR2 expression in solid tumours: a systematic review and meta-analysis. *Sci Rep*. 2017 Nov 14;7(1):15500.
75. Pizzi M, Fassan M, Realdon S, Balistreri M, Battaglia G, Giacometti C, et al. Anterior gradient 2 profiling in Barrett columnar epithelia and adenocarcinoma. *Hum Pathol*. 2012 Nov 1;43(11):1839–44.

- 889 76. DiMaio MA, Kwok S, Montgomery KD, Lowe AW, Pai RK. Immunohistochemical
890 panel for distinguishing esophageal adenocarcinoma from squamous cell
891 carcinoma: a combination of p63, cytokeratin 5/6, MUC5AC, and anterior
892 gradient homolog 2 allows optimal subtyping. *Hum Pathol.* 2012 Nov
893 1;43(11):1799–807.
- 894 77. DepMap Data Explorer [Internet]. [cited 2021 Feb 12]. Available from:
895 <https://depmap.org/portal/interactive/>
- 896 78. Lee DH, Lee Y, Ryu J, Park SG, Cho S, Lee JJ, et al. Identification of Proteins
897 Differentially Expressed in Gastric Cancer Cells with High Metastatic Potential
898 for Invasion to Lymph Nodes. *Mol Cells.* 2011 Jun 30;31(6):563–71.
- 899 79. Tsuji T, Satoyoshi R, Aiba N, Kubo T, Yanagihara K, Maeda D, et al. Agr2
900 Mediates Paracrine Effects on Stromal Fibroblasts That Promote Invasion by
901 Gastric Signet-Ring Carcinoma Cells. *Cancer Res.* 2015 Jan 15;75(2):356–66.
- 902 80. Chen S, Dong H, Yang S, Guo H. Cathepsins in digestive cancers. *Oncotarget.*
903 2017 Mar 29;8(25):41690–700.
- 904 81. Bai Z, Ye Y, Liang B, Xu F, Zhang H, Zhang Y, et al. Proteomics-based
905 identification of a group of apoptosis-related proteins and biomarkers in gastric
906 cancer. *Int J Oncol.* 2011 Feb 1;38(2):375–83.
- 907 82. Wu J, Chen X hua, Wang X qiong, Yu Y, Ren J min, Xiao Y, et al. ERp19
908 contributes to tumorigenicity in human gastric cancer by promoting cell growth,
909 migration and invasion. *Oncotarget.* 2015 May 20;6(14):11794–805.

- 910 83. Jeon Y, Kwon SM, Rhee H, Yoo JE, Chung T, Woo HG, et al. Molecular and
911 radiopathologic spectrum between HCC and intrahepatic cholangiocarcinoma.
912 *Hepatology* Baltim Md. 2022 Feb 5;
- 913 84. Yosudjai J, Inpad C, Chomwong S, Dana P, Sawanyawisuth K, Phimsen S, et al.
914 An aberrantly spliced isoform of anterior gradient-2, AGR2vH promotes
915 migration and invasion of cholangiocarcinoma cell. *Biomed Pharmacother*
916 *Biomedecine Pharmacother*. 2018 Nov;107:109–16.
- 917 85. Brychtova V, Zampachova V, Hrstka R, Fabian P, Novak J, Hermanova M, et al.
918 Differential expression of anterior gradient protein 3 in intrahepatic
919 cholangiocarcinoma and hepatocellular carcinoma. *Exp Mol Pathol*. 2014
920 Jun;96(3):375–81.
- 921 86. Simon EP, Freije CA, Farber BA, Lalazar G, Darcy DG, Honeyman JN, et al.
922 Transcriptomic characterization of fibrolamellar hepatocellular carcinoma. *Proc*
923 *Natl Acad Sci U S A*. 2015 Nov 3;112(44):E5916-5925.
- 924 87. Vivekanandan P, Micchelli STL, Torbenson M. Anterior gradient-2 is
925 overexpressed by fibrolamellar carcinomas. *Hum Pathol*. 2009 Mar;40(3):293–9.
- 926 88. Yuan K, Xie K, Lan T, Xu L, Chen X, Li X, et al. TXNDC12 promotes EMT and
927 metastasis of hepatocellular carcinoma cells via activation of β -catenin. *Cell*
928 *Death Differ*. 2020 Apr;27(4):1355–68.
- 929 89. Mizuuchi Y, Aishima S, Ohuchida K, Shindo K, Fujino M, Hattori M, et al.
930 Anterior gradient 2 downregulation in a subset of pancreatic ductal
931 adenocarcinoma is a prognostic factor indicative of epithelial–mesenchymal
932 transition. *Lab Invest*. 2015 Feb;95(2):193–206.

- 933 90. Riener MO, Pilarsky C, Gerhardt J, Grützmann R, Fritzsche FR, Bahra M, et al.
934 Prognostic significance of AGR2 in pancreatic ductal adenocarcinoma. *Histol*
935 *Histopathol.* 2009 Sep;24(9):1121–8.
- 936 91. Brychtova V, Hermanova M, Karasek P, Lenz J, Selingerova I, Vojtesek B, et al.
937 Anterior gradient 2 and mucin 4 expression mirrors tumor cell differentiation in
938 pancreatic adenocarcinomas, but aberrant anterior gradient 2 expression
939 predicts worse patient outcome in poorly differentiated tumors. *Pancreas.* 2014
940 Jan;43(1):75–81.
- 941 92. Ramachandran V, Arumugam T, Wang H, Logsdon CD. Anterior Gradient 2 Is
942 Expressed and Secreted during the Development of Pancreatic Cancer and
943 Promotes Cancer Cell Survival. *Cancer Res.* 2008 Oct 1;68(19):7811–8.
- 944 93. Arumugam T, Deng D, Bover L, Wang H, Logsdon CD, Ramachandran V. New
945 Blocking Antibodies against Novel AGR2-C4.4A Pathway Reduce Growth and
946 Metastasis of Pancreatic Tumors and Increase Survival in Mice. *Mol Cancer*
947 *Ther.* 2015 Apr;14(4):941–51.
- 948 94. Kim HS, Kang SH, Park CH, Yang WI, Jeung HC, Chung HC, et al. Genome-
949 wide molecular characterization of mucinous colorectal adenocarcinoma using
950 cDNA microarray analysis. *Oncol Rep.* 2011 Mar 1;25(3):717–27.
- 951 95. Valladares-Ayerbes M, Blanco-Calvo M, Reboredo M, Lorenzo-Patiño MJ,
952 Iglesias-Díaz P, Haz M, et al. Evaluation of the adenocarcinoma-associated
953 gene AGR2 and the intestinal stem cell marker LGR5 as biomarkers in
954 colorectal cancer. *Int J Mol Sci.* 2012;13(4):4367–87.

96. Dahal Lamichane B, Jung SY, Yun J, Kang S, Kim DY, Lamichane S, et al. AGR2 is a target of canonical Wnt/ β -catenin signaling and is important for stemness maintenance in colorectal cancer stem cells. *Biochem Biophys Res Commun*. 2019 Aug 6;515(4):600–6.
97. Ma SR, Wang WM, Huang CF, Zhang WF, Sun ZJ. Anterior gradient protein 2 expression in high grade head and neck squamous cell carcinoma correlated with cancer stem cell and epithelial mesenchymal transition. *Oncotarget*. 2015 Apr 20;6(11):8807–21.
98. Chi J, Zhang H, Hu J, Song Y, Li J, Wang L, et al. AGR3 promotes the stemness of colorectal cancer via modulating Wnt/ β -catenin signalling. *Cell Signal*. 2020 Jan;65:109419.
99. Al-Shaibi AA, Abdel-Motal UM, Hubrack SZ, Bullock AN, Al-Marri AA, Agrebi N, et al. Human AGR2 Deficiency Causes Mucus Barrier Dysfunction and Infantile Inflammatory Bowel Disease. *Cell Mol Gastroenterol Hepatol*. 2021;12(5):1809–30.
100. Zheng W, Rosenstiel P, Huse K, Sina C, Valentonyte R, Mah N, et al. Evaluation of AGR2 and AGR3 as candidate genes for inflammatory bowel disease. *Genes Immun*. 2006 Jan;7(1):11–8.
101. Camarillo GF, Goyon EI, Zuñiga RB, Salas LAS, Escárcega AEP, Yamamoto-Furusho JK. Gene Expression Profiling of Mediators Associated with the Inflammatory Pathways in the Intestinal Tissue from Patients with Ulcerative Colitis. *Mediators Inflamm*. 2020;2020:9238970.

- 977 102. Jess T, Frisch M, Simonsen J. Trends in overall and cause-specific mortality
 978 among patients with inflammatory bowel disease from 1982 to 2010. Clin
 979 Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc. 2013
 980 Jan;11(1):43–8.
- 981 103. Olén O, Erichsen R, Sachs MC, Pedersen L, Halfvarson J, Askling J, et al.
 982 Colorectal cancer in Crohn's disease: a Scandinavian population-based cohort
 983 study. Lancet Gastroenterol Hepatol. 2020 May;5(5):475–84.
- 984 104. Velcich A, Yang W, Heyer J, Fragale A, Nicholas C, Viani S, et al. Colorectal
 985 cancer in mice genetically deficient in the mucin Muc2. Science. 2002 Mar
 986 1;295(5560):1726–9.
- 987 105. Tiemann K, Garri C, Lee SB, Malihi PD, Park M, Alvarez RM, et al. Loss of ER
 988 retention motif of AGR2 can impact mTORC signaling and promote cancer
 989 metastasis. Oncogene. 2019 Apr;38(16):3003–18.
- 990 106. Worfolk JC, Bell S, Simpson LD, Carne NA, Francis SL, Engelbertsen V, et al.
 991 Elucidation of the AGR2 Interactome in Esophageal Adenocarcinoma Cells
 992 Identifies a Redox-Sensitive Chaperone Hub for the Quality Control of MUC-
 993 5AC. Antioxid Redox Signal. 2019 Nov 20;31(15):1117–32.
- 994 107. Guo H, Chen H, Zhu Q, Yu X, Rong R, Merugu SB, et al. A humanized
 995 monoclonal antibody targeting secreted anterior gradient 2 effectively inhibits the
 996 xenograft tumor growth. Biochem Biophys Res Commun. 2016 Jun
 997 17;475(1):57–63.
- 998 108. Vanderlaag KE, Hudak S, Bald L, Fayadat-Dilman L, Sathe M, Grein J, et al.
 999 Anterior gradient-2 plays a critical role in breast cancer cell growth and survival

1000 by modulating cyclin D1, estrogen receptor- α and survivin. Breast Cancer Res.
1001 2010 Jun 4;12(3):R32.

1002 109. Guo H, Zhu Q, Yu X, Merugu SB, Mangukiya HB, Smith N, et al. Tumor-
1003 secreted anterior gradient-2 binds to VEGF and FGF2 and enhances their
1004 activities by promoting their homodimerization. Oncogene. 2017 Sep
1005 7;36(36):5098–109.

1006 110. Jia M, Guo Y, Zhu D, Zhang N, Li L, Jiang J, et al. Pro-metastatic activity of
1007 AGR2 interrupts angiogenesis target bevacizumab efficiency via direct
1008 interaction with VEGFA and activation of NF- κ B pathway. Biochim Biophys Acta
1009 BBA - Mol Basis Dis. 2018 May 1;1864(5, Part A):1622–33.

1010 111. Negi H, Merugu SB, Mangukiya HB, Li Z, Zhou B, Sehar Q, et al. Anterior
1011 Gradient-2 monoclonal antibody inhibits lung cancer growth and metastasis by
1012 upregulating p53 pathway and without exerting any toxicological effects: A
1013 preclinical study. Cancer Lett. 2019 May 1;449:125–34.

1014 112. Liu Z, Zhang Y, Niu Y, Li K, Liu X, Chen H, et al. A systematic review and meta-
1015 analysis of diagnostic and prognostic serum biomarkers of colorectal cancer.
1016 PloS One. 2014;9(8):e103910.

1017 113. Lee HJ, Hong CY, Kim MH, Lee YK, Nguyen-Pham TN, Park BC, et al. In vitro
1018 induction of anterior gradient-2-specific cytotoxic T lymphocytes by dendritic
1019 cells transduced with recombinant adenoviruses as a potential therapy for
1020 colorectal cancer. Exp Mol Med. 2012 Jan;44(1):60–7.

1021 114. Makawita S, Dimitromanolakis A, Soosaipillai A, Soleas I, Chan A, Gallinger S,
1022 et al. Validation of four candidate pancreatic cancer serological biomarkers that
1023 improve the performance of CA19.9. *BMC Cancer*. 2013 Sep 3;13:404.

1024 115. Białobrzaska W, Dziąbowska K, Lisowska M, Mohtar MA, Muller P, Wojtesek B,
1025 et al. An Ultrasensitive Biosensor for Detection of Femtogram Levels of the
1026 Cancer Antigen AGR2 Using Monoclonal Antibody Modified Screen-Printed Gold
1027 Electrodes. *Biosensors*. 2021 Jun 7;11(6):184.

1028 116. Arshavsky-Graham S, Ward SJ, Massad-Ivanir N, Scheper T, Weiss SM, Segal
1029 E. Porous Silicon-Based Aptasensors: Toward Cancer Protein Biomarker
1030 Detection. *ACS Meas Sci Au*. 2021 Oct 20;1(2):82–94.

1031 117. Ye X, Wu J, Li J, Wang H. Anterior Gradient Protein 2 Promotes Mucosal Repair
1032 in Pediatric Ulcerative Colitis. *BioMed Res Int*. 2021;2021:6483860.

1033 118. Vesiclepedia: Gene summary [Internet]. [cited 2022 Jul 26]. Available from:
1034 http://microvesicles.org/gene_summary?gene_id=51060

1035 119. Fessart D, Domblides C, Avril T, Eriksson LA, Begueret H, Pineau R, et al.
1036 Secretion of protein disulphide isomerase AGR2 confers tumorigenic properties.
1037 Ivaska J, editor. *eLife*. 2016 May 30;5:e13887.

1038 120. Dong A, Gupta A, Pai RK, Tun M, Lowe AW. The Human Adenocarcinoma-
1039 associated Gene, AGR2, Induces Expression of Amphiregulin through Hippo
1040 Pathway Co-activator YAP1 Activation. *J Biol Chem*. 2011 May
1041 20;286(20):18301–10.

1042 121. Jessop CE, Watkins RH, Simmons JJ, Tasab M, Bulleid NJ. Protein disulphide
1043 isomerase family members show distinct substrate specificity: P5 is targeted to
1044 BiP client proteins. *J Cell Sci.* 2009 Dec 1;122(Pt 23):4287–95.

1045 122. Gupta A, Dong A, Lowe AW. AGR2 gene function requires a unique
1046 endoplasmic reticulum localization motif. *J Biol Chem.* 2012 Feb
1047 10;287(7):4773–82.

1048

Localization	Protein	Function	References
Extracellular	AGR1	<i>Found in extracellular vesicles</i>	(118)
	AGR2	Angiogenesis/ VEGF potentialization <i>EMT Induction</i> Non canonc Wnt dependent cell migration AREG secretion and EGFR activation Binds to C4.4 and promote cell migration <i>Pro-inflammatory</i>	(71,108–110) (71,119) (71) (120) (31,93) (34)
	AGR3	Pro-migratory through SRC activation Binds to C4.4 and promote cell migration	(28) (31)
	AGR1	Activation of ATF6alpha Folding of PTX3	(42,43) (121)
	AGR2	<i>Productive protein folding and quality control</i> ERAD Regulation of PDI in cancer cells Deregulates proteins of the ubiquitin-proteasome degradation pathway CDX2 expression induction	(1,10,45) (10) (45) (45) (122)
ER	AGR3	ND	
	AGR1	ND	
	AGR2	Inhibition of p53 <i>Modulation of RuvbL2/Reptin functions</i> Stimulates cell proliferation by overexpression of cyclin D1, c-myc, p-SRC, survivin and AREG Facilitates EMT by stabilizing p65 and activation of NFkB	(44,55,56) (41) (108,120) (110)
	AGR3	ND	
	AGR1	ND	
Cytosol	AGR2	Inhibition of p53 <i>Modulation of RuvbL2/Reptin functions</i> Stimulates cell proliferation by overexpression of cyclin D1, c-myc, p-SRC, survivin and AREG Facilitates EMT by stabilizing p65 and activation of NFkB	(44,55,56) (41) (108,120) (110)
	AGR3	ND	
	AGR1	ND	
	AGR2	Inhibition of p53 <i>Modulation of RuvbL2/Reptin functions</i> Stimulates cell proliferation by overexpression of cyclin D1, c-myc, p-SRC, survivin and AREG Facilitates EMT by stabilizing p65 and activation of NFkB	(44,55,56) (41) (108,120) (110)
	AGR3	ND	

Table 1: Pro-oncogenic gains-of-functions associated with the different subcellular localizations of AGR proteins – *Italic* type setting is indicative of potential indirect pro-oncogenic functions.

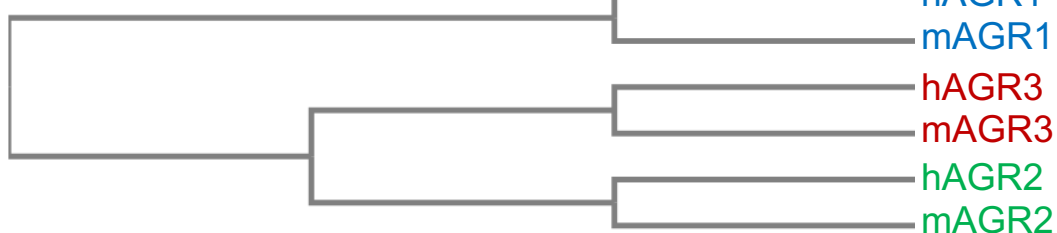
Figure legends

Figure 1 – the mammalian AGR proteins - structural aspects. **a)** Phylogram representation of human and mouse AGR1, 2, 3 proteins (performed using ClustalOmega). **b)** Amino-acid (aa) sequence alignment of human and mouse AGR1, 2 and 3 proteins. Three functional domains are indicated (Signal peptide for the targeting to the ER; thioredoxin domain; peptide binding and ER retention). Polar and basic aa are indicated in pink, polar and acidic aa in blue, polar and neutral aa in green, non-polar and neutral in red. AA identity is indicated by “*”, conservation by “:”. **c)** Structure of the E60-K64 interaction dependent AGR2 dimer. **d)** Structure of the C81 interaction dependent AGR2 dimer.

Figure 2 – Sub-cellular functions of AGR proteins: structural aspects. Demonstrated sub-cellular localizations of AGR1 (Blue), AGR2 (green) and AGR3 (red) and associated functions. ERQC: Endoplasmic Reticulum Quality Control; EV: extracellular vesicle, EMT: Epithelial-to-Mesenchymal Transition; PDI: Protein Disulfide Isomerase; UPR: Unfolded Protein Response.

Figure 3 – Expression of AGR family members in the normal gastro-intestinal tract. **a)** AGR2 expression measured by immune-histochemistry. Black borders indicate images obtained in-house and orange borders indicate data from the Human Protein Atlas (<https://www.proteinatlas.org/> - 50, 51). **b)** AGR1, AGR2 and AGR3 gene expression data extracted from the Human Protein Atlas measured by RNA-sequencing in the colon, stomach, liver and pancreas of healthy tissues.

Figure 4 – Expression of AGR family members in gastro-intestinal cancers. a)
Percentage of tumors from liver, pancreas, stomach and intestinal sphere displaying a
positive AGR2 (pink) or AGR3 (blue) immune-histochemistry staining (full circle =
100%). **b)** AGR1, AGR2 and AGR3 gene expression data extracted measured by RNA-
sequencing in the colon, stomach, liver and pancreas of cancer tissues. Data extracted
from the Human Protein Atlas (<https://www.proteinatlas.org/> - 50, 51).



b

	Signal peptide		
hAGR1	METRPRLGATCLL--GFSFLL-----LVISSDGHNGLGKGFGDHIHWRTLE	44	
mAGR1	--MSLRFGATCLL--SFSFLL-----LITSSDGTGLGKGFGDHIHWRTLE	42	
hAGR3	MMLHSALGL-CLLLVTVSSNLA-----IAIKKEKRPPQTLSRGWGDDITWVQTY	48	
mAGR3	-MLHSALAL-CLLLITVSSNLA-----IAIKKEKRPPQTLSRGWGDDITWVQTY	47	
hAGR2	-MEKIPVSA-FLLLVALSYTLARDTTVKPGAKKDKDSRPKLPQTLSRGWGDLIWTQTY	58	
mAGR2	-MEKFSVSA-ILLLVAISGTLAKDTTVKSGAKKDPKDSRPKLPQTLSRGWGDLIWTQTY	58	
	.. ** .*		*.***.:
			Thioredoxin
hAGR1	DGKKEAA-----ASGLPLMVIIHKSWCGACKALKPKFAESTEISEL	85	
mAGR1	DGKKEAAARILGLFGLGWRNPELVQADGLPLMVIIHKSWCGACKALKPKFAESTEISEL	102	
hAGR3	EEGL-----FYAQKSKKPLMVIHHLDCQYSQALKKVFAQNEEIQEM	90	
mAGR3	EEGL-----FHARKSNKPLMVIHHLDCQYCQALKKEFAKNEEIQEM	89	
hAGR2	EEAL-----YKSKTSNKPLMIIHHLDECPHSQALKKVFAENKEIQKL	100	
mAGR2	EEAL-----YRSKTSNRPLMVIHHLDECPHSQALKKVFAEHKEIQKL	100	
	:		.***:* * . *
			.*** **:
			**..:
hAGR1	S-HNFVMVNLEDEEPEKDEDFSPDGGYIPRILFLDPSGKVHPEIINENGNPYSYKYFYVSA	144	
mAGR1	S-HNFVMVNLEDEEPEKDEDFSPDGGYIPRILFLDPSGKVRPEIINESGNPYSYKYFYVSA	161	
hAGR3	AQNKFIMLNLMH--ETTDKNLSPDGQYVPRIMFVDPSTVRADIAGRYSNRLYTYEPRDL	148	
mAGR3	AQNDFIMLNLMH--ETTDKNLSPDGQYVPRIMFVDPSTVRADITGRYSNRLYTYEPQDL	147	
hAGR2	A-EQFVLLNLVY--ETTDKHLSPDGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADT	157	
mAGR2	A-EQFVLLNLVY--ETTDKHLSPDGQYVPRIVFVDPSTVRADITGRYSNRLYAYEPSDT	157	
	: ..*::*** *		*:..***** *:***:***
			.*: *: .. *
			* *
			.
			ER retention motif
hAGR1	EQVVQGMKEAQERLTGDAFRKKHLEDEL	172	
mAGR1	EQVVQGMKEAQERLTGDAFREKHFQDEL	189	
hAGR3	PLLIENMKKALRLIQSEL	166	
mAGR3	PMLVDNMKKALRLIQSEL	165	
hAGR2	ALLLDNMKKALKLLKTEL	175	
mAGR2	ALLYDNMKKALKLLKTEL	175	
	: :.***:		:

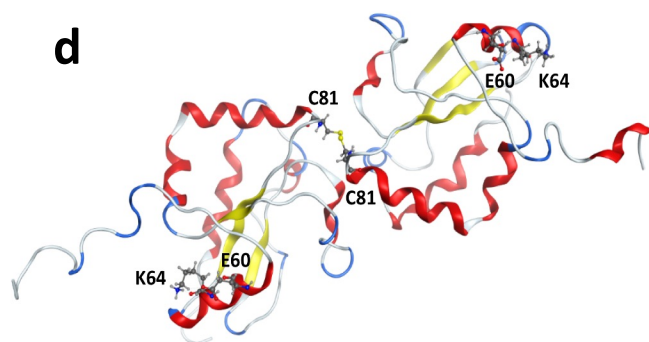
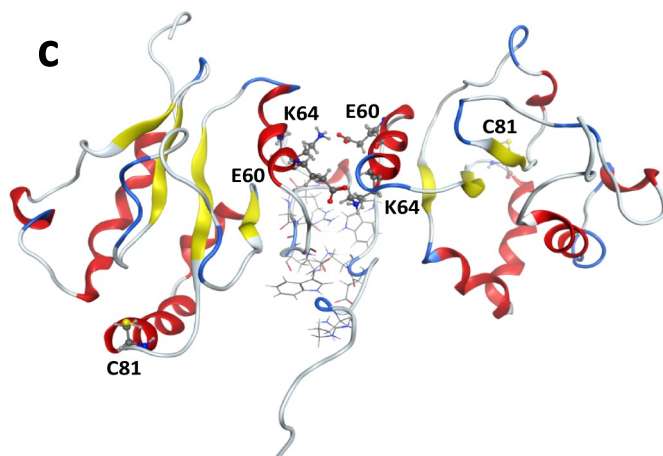


Figure 1

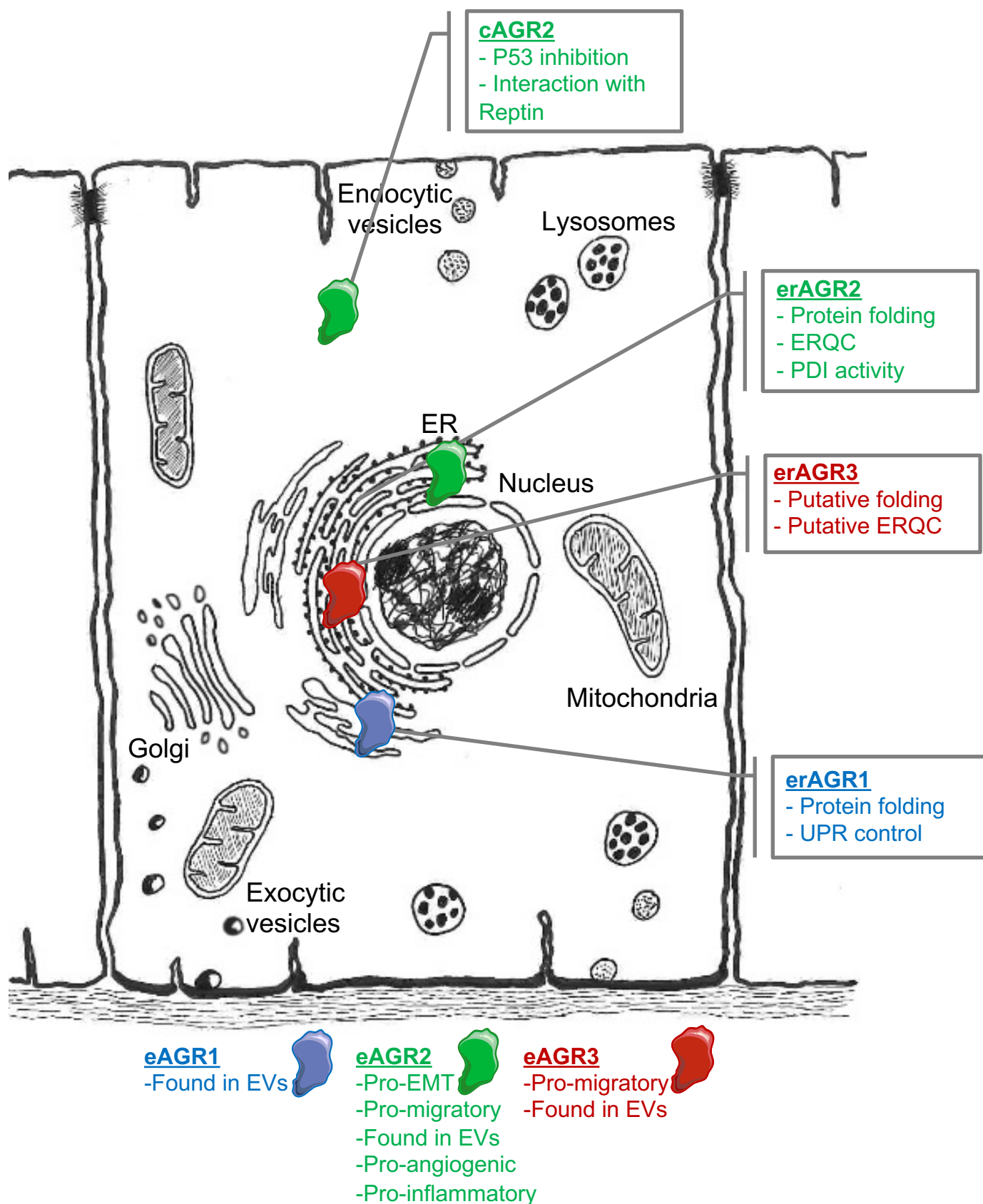
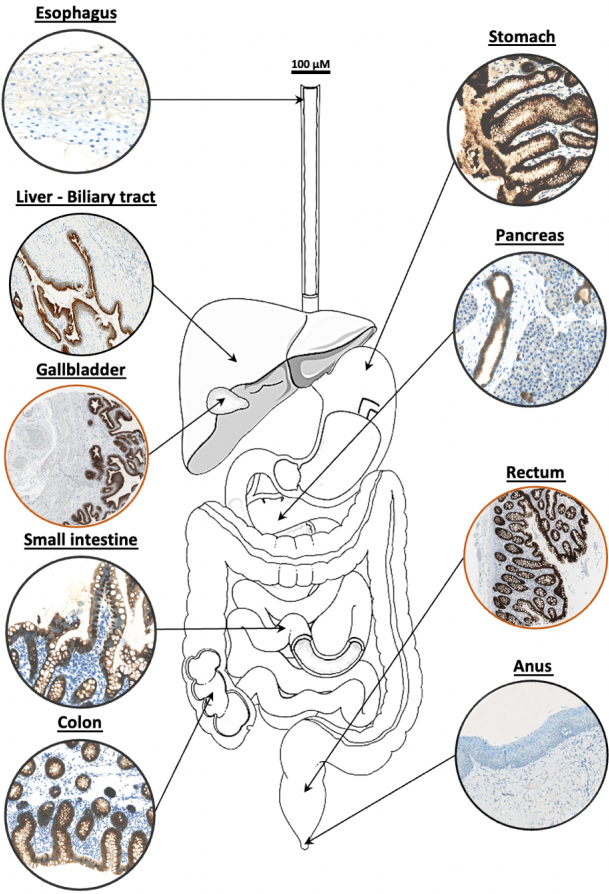


Figure 2

a



b

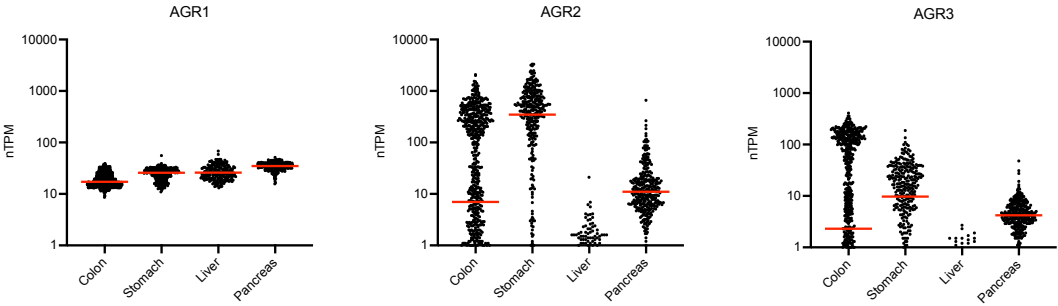


Figure 3

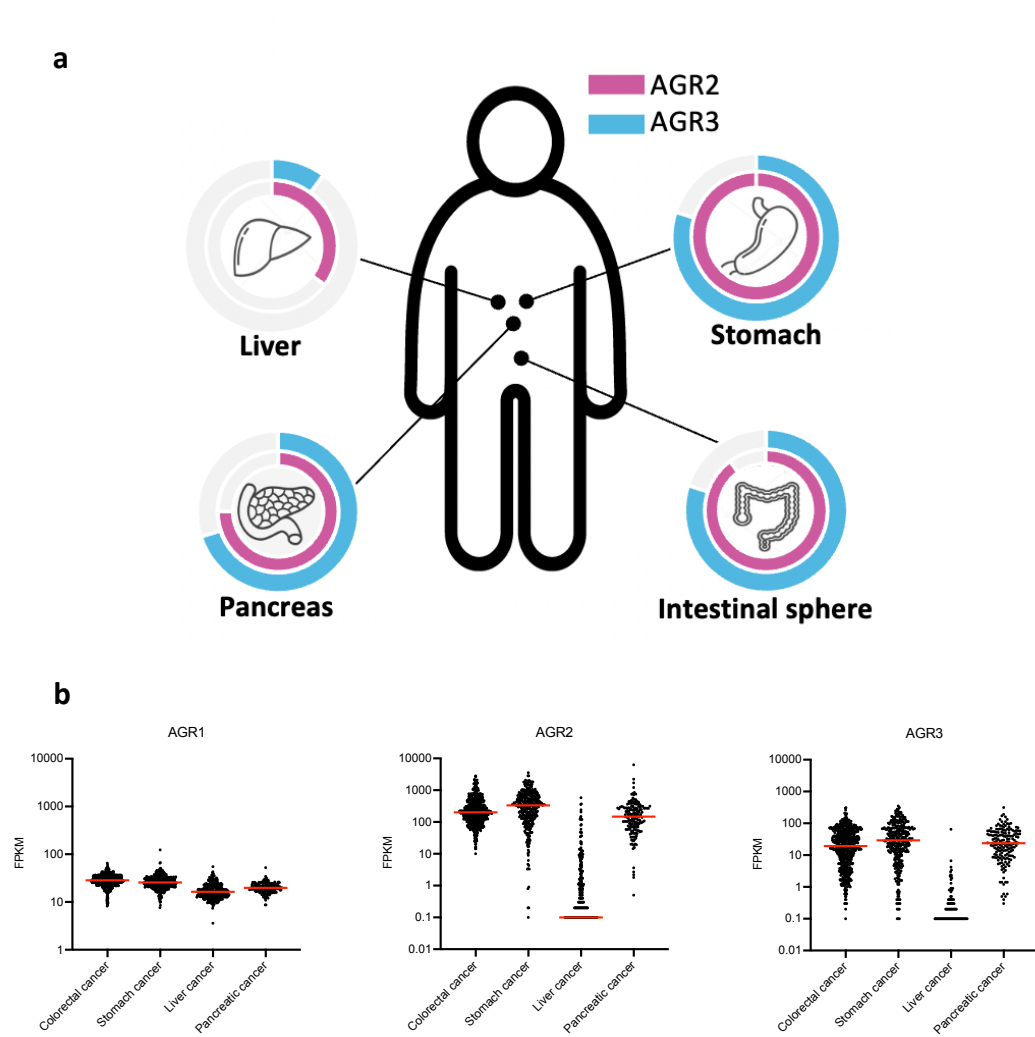


Figure 4