

Article

Not peer-reviewed version

Evaluation of the Use of Cell Lines in Studies of Selenium-Dependent Glutathione Peroxidase 2 (GPX2) Involvement in Colorectal Cancer

[R Steven Esworthy](#) *

Posted Date: 2 September 2024

doi: 10.20944/preprints202409.0125.v1

Keywords: selenium-dependent glutathione peroxidase; peroxiredoxins; NADPH oxidase 1; colorectal cancer; cell lines; public databases



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Evaluation of the Use of Cell Lines in Studies of Selenium-Dependent Glutathione Peroxidase 2 (GPX2) Involvement in Colorectal Cancer

R. Steven Esworthy

Department of Cancer Genetics and Epigenetics, Beckman Research Institute, City of Hope National Medical Center, Duarte, CA, USA, 91010; sesworthy@coh.org

Abstract: Hydroperoxides (ROOH) are known as damaging agents capable of mediating mutation, while a role as signaling agents through oxidation of protein sulfhydryls that can alter cancer-related pathways has gained traction. Glutathione peroxidase 2 (GPX2) is an antioxidant enzyme that reduces ROOH at the expense of glutathione (GSH). GPX2 is noted for a tendency of large increases or decreases in expression levels during tumorigenesis that leads to investigators focusing on its role in cancer. However, GPX2 is only one component of multiple enzyme families that metabolize ROOH and GPX2 levels are often very low in the context of these other ROOH reducing activities. Colorectal cancer (CRC) was selected as a case study for examining GPX2 function as colorectal tissues and cancers are sites where *GPX2* is highly expressed. A case can be made for a significant impact of changes in expression levels. There is also a link between GPX2 and NADPH oxidase 1 (NOX1) from earlier studies that is seldom addressed and is discussed, presenting data on a unique association in colon and CRC. Tumor-derived cell lines are quite commonly used for pre-clinical studies involving the role of GPX2 in CRC. Generally, selection for this type of work is limited to identifying cell lines based on high and low *GPX2* expression with the standard research scheme of overexpression in low expressing lines and suppression in high expressing lines to identify impacted pathways. This overlooks CRC subtypes among cell lines involving a wide range of gene expression profiles and a variety of driver mutation differences along with a large difference in *GPX2* expression levels. A trend for low and high *GPX2* expressing cell lines to segregate into different CRC subclasses, indicated in this report, suggests that choices based solely on *GPX2* levels may provide misleading and conflicting results by disregarding other properties of cell lines and failing to factor in differences in potential protein targets of ROOH. CRC and cell line classification schemes are presented here that were intended to assist workers in performing pre-clinical studies but are largely unnoted in work on GPX2 and CRC. Studies are often initiated on the premise that the transition from normal to CRC is associated with up-regulation of *GPX2*. This is probably correct. However, the source normal cells for CRC could be almost any colon cell type, some with very high *GPX2* levels. These factors are addressed in this study.

Keywords: selenium-dependent glutathione peroxidase; peroxiredoxins; NADPH oxidase 1; colorectal cancer; cell lines; public databases

1. Introduction

Selenium-dependent glutathione peroxidase 2 (GPX2) is a member of the GPX family of antioxidant enzymes that reduce peroxides such as H_2O_2 and fatty acid hydroperoxides (ROOH) at the expense of the tripeptide, glutathione (GSH)[1]. When GPX2 is expressed it is generally in the company of its sister isoenzyme, GPX1, and the peroxiredoxin (PRDX) family of ROOH reducing activities that use thioredoxin as a reducing agent. Since the initial characterization of *GPX2* gene expression in the gastrointestinal (GI) tract and other human and rodent tissues, many studies have examined the roles of GPX2 in carcinogenesis in epithelial cells [1,2]. Too often, the full context of antioxidant enzyme families is overlooked leading to claims that GPX2 is having strong effects in cases where its expression is low by comparison [1,2]. There may be a basic misconception that the ability to detect the mRNA by RT-PCR, the protein by Western blotting or by immunohistochemistry (IHC) represents significant expression levels by not examining GPX1 or PRDXs [1,2]. In several studies using RT-PCR, *GPX2* levels seemed to be at the lower limits of detection (sometimes by

comparison to cell lines with established high *GPX2* expression levels) and yet the protein was still detectable on Western blots, encouraging the view that *GPX2* could be exerting a significant impact in this lower range [1,2]. In one instance, *GPX2* was shown to be detected in glioblastoma multiform (GBM) by IHC from The Human Protein Atlas (THPA), a cancer that barely expresses *GPX2* [1]. The Human Protein Atlas warns that *GPX2* antibodies used in that analysis appear to exhibit off target binding (<https://www.proteinatlas.org/>, as of 5/2024) and declares that, generally, the protein was not detected in GBM, while in the detailed comments suggesting that Purkinje cells and not glia show some *GPX2* staining in normal samples.

GPX1, with a nearly identical hydroperoxide substrate range, was not included in most of these studies. This is unfortunate because *GPX2* expression is often minute compared to *GPX1* (GBM) or at best comparable, and the sum of both is often small compared to the ubiquitous peroxiredoxin antioxidant collective (*PRDX1-6*) [1]. Thus, *GPX2* seems to be exerting effects which are not truly explained despite the apparent variation in cellular oxidant levels after manipulation of *GPX2* levels [1]. Colon/rectum and colorectal cancer (CRC) are cases where *GPX2* levels are high enough to have an impact on ROOH levels. In the overall scheme of ROOH antioxidants, *GPX1* and *GPX2* seem to act solely as reducing agents, maintaining low peroxide tone. In this context, low peroxide tone means preventing direct oxidation by ROOH and damage to DNA, cell membranes and proteins. PRDXs are postulated to have two general functions. In competition with GPXs, PRDXs would be acting as general antioxidants. Second, PRDXs may act as oxidant relays, eliminating ROOH species as potential damaging agents and safely registering and conveying the oxidant signal as PRDX-SOH to target proteins (Figure 1) [3,4]. PRDX1 and 2 would tend to dominate ROOH metabolism by virtue of high levels of expression and initial ROOH rate constants greater than *GPX1* and *GPX2*. The mechanism of action of *GPX1* and *GPX2* at high levels of expression in this model is diversion of ROOH away from the PRDX relay, thereby nonspecifically lowering the impact of ROOH on signaling. This is a real possibility in CRC, esophageal and stomach cancers as shown in this report. Possible involvement of *GPX2* in other cancers was evaluated in ref. [1]. The downstream effects and specificity would be dependant on the general gene expression profile of the cell types. In CRC as many as six general expression profiles are noted. Thus, different downstream effects could arise from the impact of *GPX2* based on widely different general gene expression profiles and not a direct specific effect of *GPX2*. In this investigation cell lines are sorted with the finding that cell lines with low and high *GPX2* levels end up in different profile categories. While this is of interest, the very low expression in some cell lines is at odds with the expression levels of the corresponding CRC types, calling into question their value in studies. A common trope in studies on *GPX2* roles across many cancer types is overexpression in low expressing cell lines and reciprocal knockdown in high expressing cell lines [1,5]. Thus, different downstream effects could arise from the impact of *GPX2* in each category based on widely different general gene expression profiles and not a direct specific effect of *GPX2*. The goal is to provide some guidance in selecting CRC-derived cell lines to study colorectal cancers in the context of *GPX2* expression, specifically selection of certain cell lines to accomplish study of CRC types that tend to relapse and metastasize, avoiding use of cell lines with exceedingly low *GPX2* levels. The histological and functional links between superoxide generating NADPH oxidase 1 (NOX1) and *GPX2* have been demonstrated in colon/rectum, CRC and, cell lines and are discussed in the context of cell line choices. The goal is to provide some guidance in selecting CRC-derived cell lines to study colorectal cancers in the context of *GPX2* expression, specifically selection of certain cell lines to accomplish study of CRC types that tend to relapse and metastasize, avoiding use of cell lines with exceedingly low *GPX2* levels. The histological and functional links between superoxide generating NADPH oxidase 1 (NOX1) and *GPX2* have been demonstrated in colon/rectum, CRC and, cell lines and are discussed in the context of cell line choices.

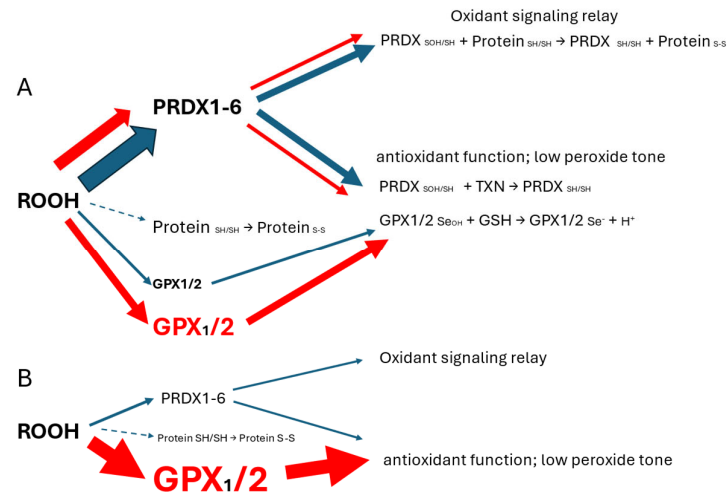


Figure 1. A. GPX1 and GPX2 are in competition for ROOH with PRDX. The greater abundance of PRDXs in most cell types and the higher rate constants of PRDX1 and PRDX2 (2-5-fold) would divert a major portion of ROOH into a signaling relay in combination with a general antioxidant function (blue arrows), largely preventing direct oxidation of proteins (dashed arrow) [3]. The reduction of oxidized PRDXs may be slow (interaction with protein targets and bottleneck at thioredoxin reductase; thioredoxin, TXN; glutathione, GSH), so that GPX1 and GPX2 take up the slack while PRDXs regenerate back to the reduced form, maintaining the required low peroxide tone [4]. When GPX2 and/or GPX1 attain higher relative levels (RED), they would siphon ROOH from the PRDX oxidant relay and alter the impact of ROOH signaling. B. In the GI tract, where a few cell types have very high levels of GPX2, there may be a dependance on GPX1/2 over and above that of repressed PRDXs to prevent direct oxidation of cell components by ROOH, demonstrable when *GPX1* and *GPX2* are knocked out in mice, leading to excess apoptosis and anoikis, the latter responsible for inflammation of the ileum and colon.

2. Materials and Methods

The primary tool used in this work was internet access to public databases, Cancer Dependency Map consortium (DepMap, <https://depmap.org/portal/>), The Human Protein Atlas/ The Cancer Genome Atlas (THPA/TCGA, <https://www.proteinatlas.org/>; <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>), Gene Expression profile Interactive Analysis (GEPIA2 <http://gepia.cancer-pku.cn/>), Tumor IMMune Estimation Resource (TIMER2.0, (<http://timer.cistrome.org/>) and PubMed (<https://pubmed.ncbi.nlm.nih.gov/>). Analysis spanned the time frame of 1/2024 to 6/2024). The use of data from these sites was previously explored and validated against results from experimental studies with some discrepancies noted [1]. The extrapolation of TPM to protein levels was examined and found to be reasonably consistent [1]. The analysis relied on basic tools accessible to most investigators. Excel was used to collate and process datasets and generate tables and some graphs; GraphPad PRISM 9.3 was used to generate graphs and perform statistical analyses; Microsoft Power Point was used to create figures. The primary analyses involved sorting CRC-derived cell lines, focusing on *GPX2* expression levels, into CRC categories based first on marker analyses in Medico et al that provided the biggest set of cell lines (150 with 65 of 77 from DepMap found in common; 5/2024) [6]. While starting with the methods of Sadanandam et al, they sorted additional cell lines with some new markers and weighting [7]. Then, they sorted the 150 cell lines based on the analyses of Marisa et al, Budinska et al, De Sousa E Melo et al, and Roepman et al [8–11]. CRC subtype analysis and corresponding cell lines were taken directly from Schlicker et al, Sadanandam et al, Berg et al, and Hu et al (mucinous adenocarcinoma markers) [7,12–14]. *GPX1*, *GPX2*, *NOX1*, *PRDX* and *CAT* expression data for cell lines was from DepMap. Because only one value is provided for each cell line, THPA's somewhat smaller list was consulted to get some sense of the inherent error (56 cell lines in

common) for *GPX2* and *NOX1*. In the text, cell line *GPX2* expression levels are presented as the average of DepMap and THPA when both are available. In the figures, only DepMap values are presented, except to correlate with THPA. Marker analysis was performed using a short-hand list from the iCMS system and the combined expression data (average, log₂ TPM+1 for markers) [15]. In addition, for HT29, markers associated with differentiation (*MUC2* and *LYZ*), *TGFβ* expression (*CCN1*, *ITGAV*, *TGFβ*) and EMT (*FN1*, *CDH2*, *VIM*) were examined [16–18]. Statistical analysis was performed for *GPX2* expression levels among CRC subtype groupings; Mann-Whitney; log₂ transformed data directly from DepMap; PRISM 9.3.

3. Results and Discussion

3.1. Correlation of *GPX2* Levels in Tumors and Cell Lines

To ascribe the importance of antioxidant enzymes that reduce ROOH, it is necessary to demonstrate that an enzyme is present in a significant quantity (Figure 1) [1,2]. The initial characterization of *GPX2* to be highly expressed in the gastrointestinal (GI) tract and undetectable in many other tissues by Northern blotting is still valid (Figure 2A) [1,2,19]. Based on results compiled by the TIMER2.0 database, in general agreement with TCGA (ref. [1]), *GPX2* is highly expressed at the tissue level in the mid-lower GI (corresponding cancer, TCGA abbr.; colon/rectum, COAD/READ; 457 COAD samples vs. 41 normal and 166 READ samples vs. 10 normal in Figure 2A; 5/2024), bladder (BLCA; 408 vs. 19 normal), esophagus (ESCA; 189 vs. 11 normal), head and neck tissues (HNSC; 520 vs. 44 normal), liver (LIHC; 371 vs. 50 normal), stomach (STAD; 415 vs. 35 normal), in cancerous pancreas (PAAD; 178 vs. 4 normal; median for normal is several logs higher than other database values; ref. [1]) and lung squamous cell carcinoma (LUSC; 501 vs. 51 normal), although at a lower level in lung adenocarcinoma (LUAD; 515 vs. 59 normal) (Figure 2A) [1]. High expression in normal tissues is often confined to a few cell types. The small intestine is a prime example of the limited zone of expression within a high expressing tissue, apparently just the easily recognizable Paneth cells (Figure 3; THPA). This creates one of the problems in analysing *GPX2* using tissue level metrics from databases as will be elaborated and is not unique to the GI-tract (Figures 2 and 3) [1].

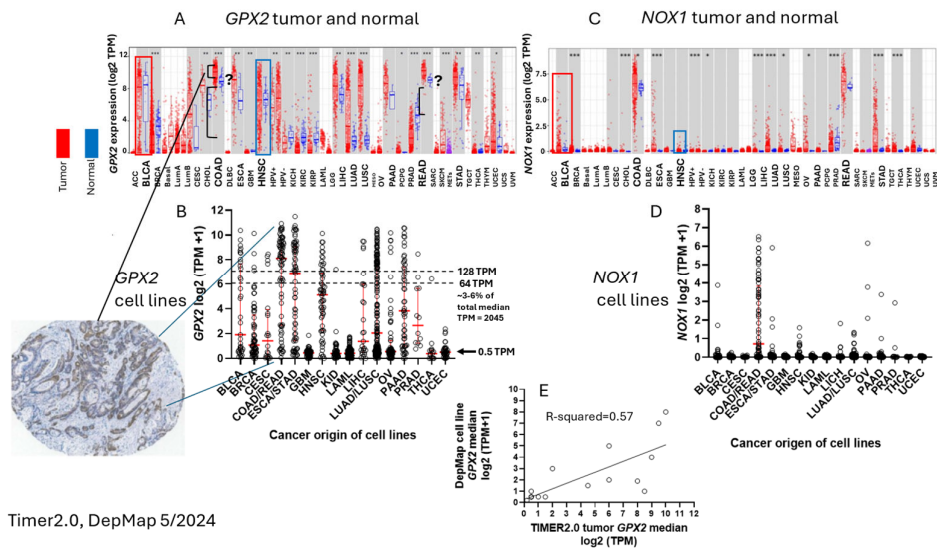


Figure 2. Comparison of normal tissue and tumor sample *GPX2* (panel A) and *NOX1* expression levels (panel C) (log₂ TPM) with levels in cancer-derived cell lines (panels B and D; log₂ TPM+1) (TIMER2.0 and DepMap-5/11/24). The CRC tumor sample *GPX2* IHC insert (THPA) is estimated to have *GPX2* expression levels falling inside the upper bracket (COAD tumors). The lower bracket denotes rare samples with very low expression. Low expressing READ tumors are also bracketed. ? denotes high *GPX2* expression from a few normal cells in colon. The dashed lines across the cell line *GPX2* expression panel denotes 2 thresholds for significant levels, relative to total antioxidant gene expression- *GPX1*, *GPX2*, *PRDXs* 1-6 and catalase- in TPM (2045 total TPM). The median *GPX2*

expression level for all cell lines is 0.5 TPM. BLCA and H&N tumor samples have broad ranges of *GPX2* expression (red and blue rectangles) along with LIHC, suggesting nothing anomalous with isolation of both high and low *GPX2* expressing cell lines, unlike COAD/READ. Panel E shows the correlation between the median tumor levels and the median cell lines levels for each tumor type. TCGA tumor abbreviation convention is used (<https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/tcga-study-abbreviations>).

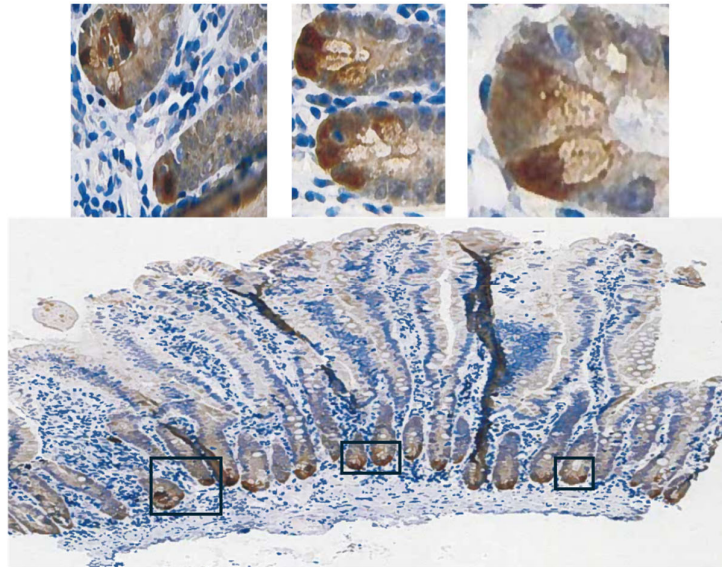


Figure 3. The few cells with high *GPX2* protein levels in small intestine are at the crypt base (lower panel) and appear to be Paneth cells (upper panels). This localized high *GPX2* expression may supersede *PRDXs* for overall antioxidant protection (Figure 1B). Some low-level expression may extend to the top of the crypt to include other cell types (THPA: NOS (M-00100) Patient id: 2411; 5/5/2024).

Cell line *GPX2* gene expression levels were compiled on the basis of tissue and cancer of origin (DepMap; 5/2024). As shown in Figure 2B, cell lines derived from COAD/READ as well as from ESCA/STAD have a high median level of *GPX2* gene expression, while many cell lines from HNSC and PAAD also have somewhat elevated *GPX2* expression, based on median levels. There is some correlation between the respective median tumor levels and median cell lines levels (Figure 2E). The low *GPX2* levels in cell lines derived from the high expressing cancer sources, COAD/READ and ESCA/STAD, are somewhat unique among the antioxidant enzyme genes, *GPX1* and *PRDX1-6*. A few cell lines have lower than expected *GPX1* and *PRDX* levels, but nothing like *GPX2* (*GPX2*, 59/128 low expressing lines (≤ 64 TPM; ref. [1]); *GPX1*, 13/128; *PRDX1*, 0/128; *PRDX2*, 6/128; *PRDX3*, 3/128; *PRDX4*, 6/128; *PRDX5*, 0/128; *PRDX6*, 2/128; Figure 4) [1]. The low expression of *GPX2* in cell lines could be anticipated for BLCA, HNSC, LIHC, and PAAD from the variation observed in tumors and is not explained by variation among COAD/READ and ESCA/STAD samples at the tissue level (Figure 2A).

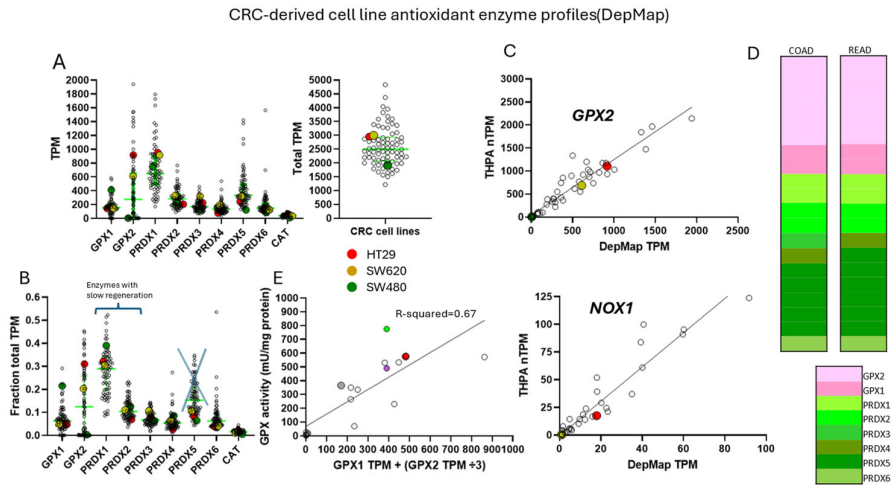


Figure 4. Antioxidant gene expression levels in CRC-derived cell lines (panels A and B; DepMap) with total levels (panel A). Error bars are median levels and quartiles. The Xed out PRDX5 set indicates the low ROOH rate constant that basically factors PRDX5 from consideration in apprising impact (panel B). Slow reduction of oxidized PRDXs 1 and 2 may bring the enzymes back down into the overall ROOH rate of metabolism with GPX1 and GPX2, although their initial ROOH rate constants are greater than GPX1 and GPX2 [59–61]. Panel C is an indication of the relative error in DepMap TPM values for GPX2 and NOX1 based on comparison with THPA nTPM as a correlation. The colorectal tumor panel shows the relative fraction of GPXs and PRDXs to total antioxidant mRNA expression in COAD and READ (panel D). The oversized dots in the cell line panels A, B, and C represent 3 cell lines discussed in the main text, HT29 (red), SW480, (dark green) and SW620 (yellow). Panel E is a comparison of GPX1/2 activity levels in cell lines determined by the author in prior work [1,2] with DepMap TPM using the formula, GPX1 TPM + (GPX2 TPM*3) [1], including CRC-derived cell lines, HT29 and CACO2, red and gray dots. The intestinal cancer cell line, HUTU80, also assayed by the author, is at 1.3 TPM and 27 mU/mg. The green and purple dots are two activity measurements of OVCAR8 (NCI/ADR-RES) separated by four years.

For a variety of reasons, the relative TPM levels among tissues and cell lines presented represent only a rough guide to thinking about where GPX2 might have a significant role (Figures 2 and 4). DepMap TPM results have some predictive value for GPX activity in cell lines (Figure 4 E; selenium-supplemented culture media for cell line activity) [1]. GPX activity at and above 190 mU/mg is in the range of normal tissue and cancer sample levels [1]. One word of caution is made with regards to the GPX TPM data of DepMap and THPA and for work in almost all current papers [1]. The two database projects and almost all current published work did not include supplementation of the culture media with selenium (10% FBS/FCS; fetal bovine/calf serum; some rare exceptions in DepMap in combination with low serum for culture) [20]. As a rule, this is suboptimal for GPX1 and GPX2 protein and activity levels in cell lines [1,2,21]. Different batches of serum can have widely different selenium levels, and this will have an impact on the protein and activity levels of both GPX1 and GPX2 and possibly on the mRNA levels of GPX1 [21–23]. GPX1 mRNA levels may be underrepresented in the databases and in studies [1]. Selenium supplementation could conceivably shift up to 8 of the 13 low expressing GPX1 level GI tract-derived cell lines into the high expression range (2-fold increase; see HepG2 in ref. [22]). In studies, the protein and GPX activities will often be up to one-half to one-fourth the optimal levels and can be even less, the lower activity a feature of cell lines expressing only GPX1 [22,23]. This could impact the reproducibility of findings and downplay the role of GPX1 in favor of GPX2 [1].

3.2. The Choice of CRC-Derived Cell Lines for a Case Study: GPX2 and NOX1

It was reported that the superoxide generating NADPH oxidase, NOX1, is a major source of oxidants that produce ileocolitis in mice deficient in both *GPX1* and *GPX2* gene expression [24,25]. While mice deficient in both *GPX1* and *GPX2* (i.e. GPX1/2-DKO; double knockout) have spontaneous ileocolitis (Figure 1B; supporting a lesser role for PRDXs in general antioxidant function), after knocking out *NOX1*, the triple KO mice no longer have gut pathology. Eliminating the H₂O₂ generating enzyme, DUOX2 (via a DUOXA- KO), relieved crypt/gland base anoikis and inflammation while having no impact on excess crypt/gland base apoptosis observed in the DKO mice. *NOX1* mRNA is uniquely, highly expressed in normal and cancerous colorectal tissues (Figure 2C). The same pattern of *NOX1* expression is observed in cell lines derived from colorectal cancerous tissues (Figure 2D). It is likely that the indicated colocalization of *GPX2* and *NOX1* in the crypt/gland epithelium (although not necessarily the same cells) of the intestine and colon allows them to keep balance in the redox status [26–28].

Some cells in the colon epithelium have naturally high expression of *GPX2* and this seems to carry over into the cancers [29–31]. The estimated levels of *GPX2* at the mRNA and protein levels in CRC show that they likely exceed that of *GPX1* and rival that of PRDX1 and PRDX2, meaning *GPX2* could have a major impact on ROOH levels (Figure 4D) [1]. This and the apparent link to *NOX1* led to choosing CRC-derived lines as the study case for trying to understand what to consider in selecting lines for cancer studies. The complexity of issues in CRC, the peculiarities of cell line properties and culture methods make this a difficult case study. While this analysis may not completely succeed, the details of the process may be instructive.

3.3. Remaining Uncertainties in GPX2 Expression in Normal Colon/Rectum

The reasons for lack of a clear answer to the choice of cell lines include no direct link between the cell types that naturally express *GPX2* at high levels and its expression in tumors. Despite decades of characterization, the exact cell expression profile in the colon is unclear [29]. With studies using single-cell profiling of normal colon and CRC available, it should be a simple matter to establish the identity of the normal high *GPX2* expressing cells and possibly infer *GPX2* expression levels for sources of CRC cells, this based on data from polyps representing an early point in the malignancy continuum [15,32]. Variable co-expression profiles between *GPX2* and *NOX1* might aid in this process (Figures 2 and 5).

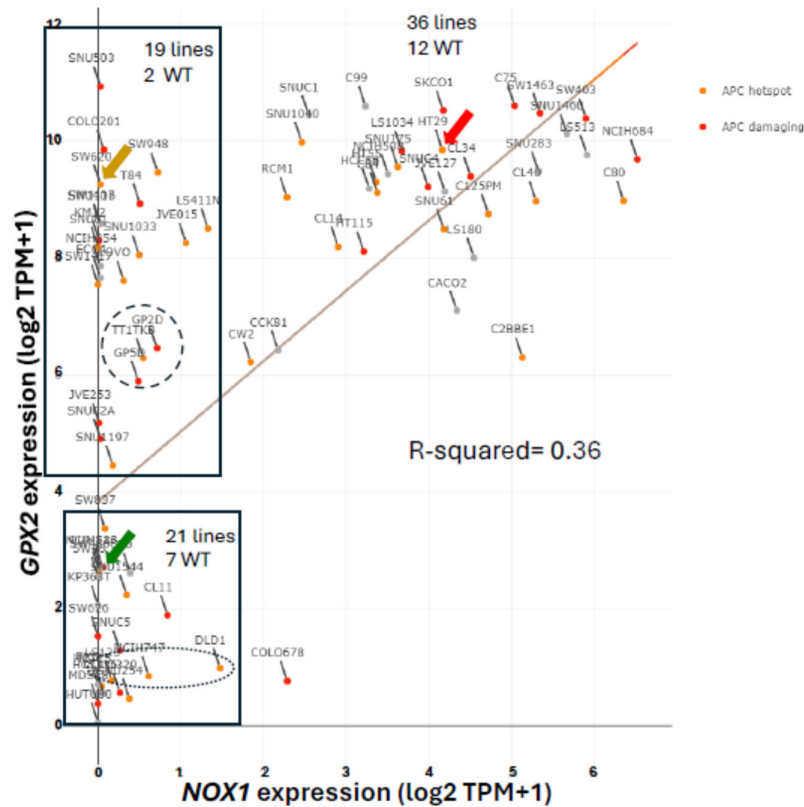


Figure 5. Correlation of *GPX2* and *NOX1* expression in the set of 77 CRC-derived cell lines in DepMap (5/2024), denoted by *APC* gene mutation status. Cell lines HT29, SW480 and SW620 are indicated with arrows. The dashed circle encloses paired cell lines, GP2D and GP5D (same primary), and the dashed oval the suspected DLD1 and HCT15 pair. WT-wildtype *APC* gene.

The possible candidates for high *GPX2* expression are Paneth cells (Figure 3; mice develop tumors in the small intestine) or the colon equivalent, deep secretory cells (so far, identified in mice and not documented for *GPX2* expression) [33]. Refined localization of *GPX2* expressing cells in human colon by IHC shows that they represent only a few cells at the base of each gland, paralleling small intestine (Figure 3) [29]. A single-cell analysis did confirm high *GPX2* expression in Paneth cells, while being somewhat unclear about other cell types, particularly in colon [34]. Paneth cells are not known to give rise to tumors, although they exhibit plasticity during loss of LGR5+ crypt cells that allows them to gain stem cell properties and repopulate crypts [35–38]. Whether they retained *GPX2* expression in this process was not determined. It is this type of plasticity exhibited across many cell types in the colon during carcinogenesis that lends itself to uncertainty over whether *GPX2* is being up regulated and to what extent [36]. Up regulation would be the default when any cell types outside of the lower crypt/glands are the source of the tumors (Figure 3) [39]. There is some evidence that up regulation of *GPX2* mRNA levels on the order of 4-5-fold in Stem-like cells occurs during the progress from early lesions to CRC [32]. Similarly, looking at whatever cells are expressing *GPX2* protein at high levels in normal colon suggests possible elevation up to 5-fold in protein levels in CRC cells (statistically significant) [29]. However, that determination did not establish any link between the normal cells and the cell types that gave rise to the tumors. If the candidate cells are not the naturally high *GPX2* expressing cells, the magnitude of elevation could be tens-hundred-fold, providing a strong rationale for *GPX2* impacting tumorigenesis by modulation of ROOH levels. There is ample evidence that an alteration of *GPX2* levels from essentially no expression to levels characteristic of most CRC samples (~27% of total antioxidant enzyme TPM; GPXs, PRDXs and catalase) would impact tumor pathways via known redox sensitive proteins, such as PTEN (Figure 1A and Figure 4D) [1,40,41].

3.4. The Problems with Databases Combined with the Issues of Normal Cell Expression

Note, the up regulation of *GPX2* expression is not being based on data from sites like TCGA/THPA, TIMER2.0 and GEPIA2. GEPIA2 data present real problems with exceptionally low *GPX2* expression found in many normal samples (Figure 6). This extends to several other genes whose expression is confined to the mucosa as opposed to those expressed in the mucosa and muscularis, like villin vs. β -actin (Figure 6).

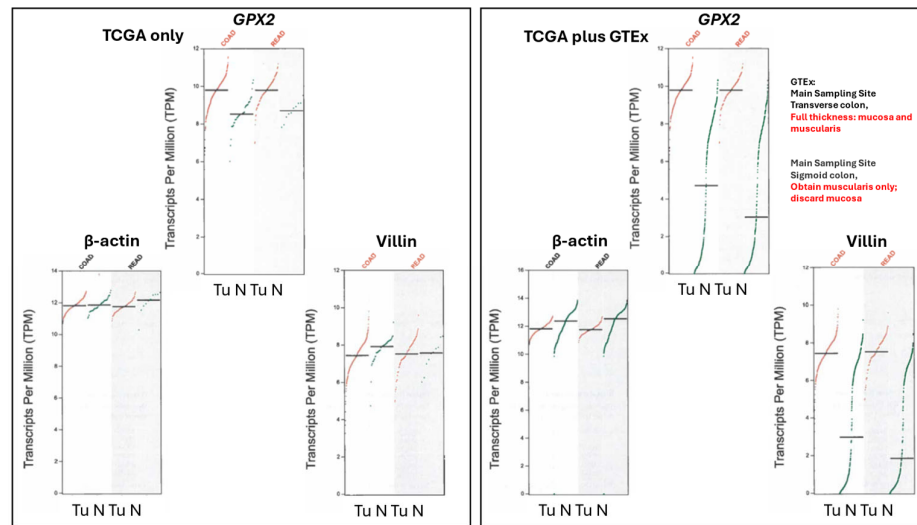


Figure 6. Discrepancies in the *GPX2* expression levels of normal samples in GEPIA based on database sources (TCGA alone or including GTEx). Red text indicates the 2 discrepant sample processing protocols mentioned in GTEx that might contribute to the outcomes. Other genes, such as *VIL* (villin) are apparently impacted by the alternative GTEx normal tissue sampling protocols; compared to β -actin.

TCGA/THPA, TIMER2.0 and GEPIA2 tumor databases are consistent in showing CRC tumors as having high *GPX2* expression, with only few exceptions (28/597 CRC samples TPM < 256, THPA; TIMER2.0 has ~18/623; Figure 2A; GEPIA2 ~12/397; Figure 6). The issue with normal samples in GEPIA2 might stem from the divergent protocols for processing between the 2 sampling sites mentioned in GTEx (<https://gtexportal.org/home/>), the source of much of the data. For transverse colon, full-thickness samples were analyzed, while for sigmoid colon, only the muscularis was sampled. *GPX2* is not expressed outside of the epithelium of the mucosa, except for scattered cancer-associated fibroblasts (CAF) (Figure 3) [29–31,42]. It is not clear that this is a complete explanation for the discrepancy between the GEPIA2 data sets and TCGA/THPA and TIMER2.0. The rectum set does not mention similarly divergent processing protocols and the sample numbers for the alternatively processed colon samples do not add up to the total number of samples listed. TCGA/THPA and TIMER2.0 show much lesser range of increase in *GPX2* expression between normal and CRC, on the order of 2-20-fold (see matched sets, ref. 43) with a few matched samples showing a decrease in CRC or no alteration [43].

Part of the issue in comparing normal to CRC using TPM/FPKM or standard IHC presentation (combined intensity and proportion scores), beyond lack of clear identification of the source cells, is the limited numbers of normal high expressing cells and the evident increase in the proportion of expressing tumor cells in CRC samples (tumor purity) (Figure 3 and 7) [29,30,44].

The upper range of the increase in tissue expression of 20-fold could be accounted for by the 4-5-fold increase in cell level expression and a 4-5-fold increase in the relative volume of the expressing cells in tumor samples, although many CRC samples seem to exceed this increase in proportion based on IHC at THPA (Figure 7). None of this might be relevant if the source tumor cells are not represented by the naturally high *GPX2* expressing normal cells. The question of source cells is complicated by alterations in the *GPX2* protein and mRNA expression pattern observed in mouse

colon after dextran sodium sulfate-induced injury. Here, the zone of high expression was expanded to the luminal region showing that many cell types in the colon can produce high GPX2 levels under stress [45]. There may be differences between mice and humans in the range of GPX2 protein expression in ileum and colon. IHC of mice seems to show high levels of protein at the crypt/gland base and moderate levels of protein detected up to the villus tip/lumen, while one IHC study of human samples seems to show expression only in the gland base and in a second the image is in gray-tone and it is unclear by inspection whether any protein was detected outside of the narrow range of high levels at the gland base [29,30,46]. The authors comment that GPX2 was detected in the ileum crypts and colon gland base and don't mention detection outside of those zones [30]. In a third case, GPX2 was detected at the lumen with decreased levels from the colon gland base, while confined to Paneth cells in the ileum [31]. Enlargement of the image of human small intestine GPX2 IHC in Figure 3 hints at very low expression in patches in the upper crypt and up to the villus tip, however, the call by THPA is that cell types other than the Paneth cells are negative for expression.

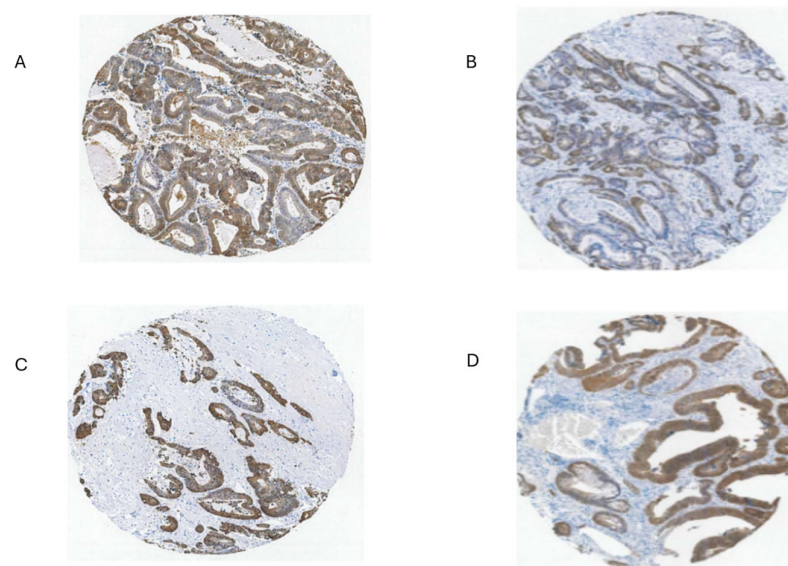


Figure 7. A set of GPX2 CRC IHC-stained samples showing some variation in staining intensity within the tumors (THPA, 2-5/2024).

3.5. Efforts to Assist Investigators in Selection of Cell Lines

The suitable selection of cell lines for CRC studies is made possible by analysis of CRC for driver and passenger mutations (classical analysis) and later efforts to classify tumors by small- or large-scale analysis of mRNA expression profiles and IHC, often combined with mutation profiles and epigenetic modification profiles; Cimp, CpG island methylator phenotype; MSI, microsatellite instability (CMS, CCS, CRIS, iCMS and others) [7–15,47–49]. In some cases, cell lines were similarly classified in attempts to match them with CRC types (Figure 8) [6–15].

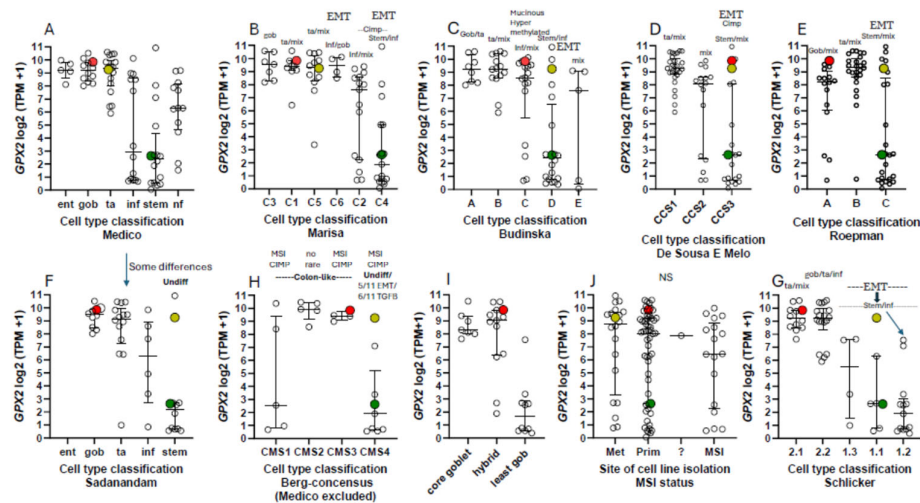


Figure 8. Classification of cell lines in various schemes into CRC subtypes with *GPX2* expression levels ($\log_2 \text{TPM}+1$) (see references 6-15). Medico et al [6] has the largest set of lines, although only 65 are found in common with DepMap. NF in panel A indicates cell lines in DepMap and not in Medico et al. Medico et al examines cell lines, exclusively, and expands on the classification scheme of Sadanandam et al [7] with some differences arising between the two with no real impact on the grouping of low and high expressing lines. The Medico/Sadanandam group names are carried over into the other schemes (panels B-H; ent-enterocyte-like; gob-goblet-like; ta-transit amplifying-like; inf-inflammatory-like; stem-stem cell-like; nf-not found in Medico et al) since the information was obtained from processing by Medico et al, except for panels F, H and G. Properties of the groups are listed above in panels B-H (EMT-epithelial to mesenchymal transition; Cimp- CpG island methylator phenotype; MSI-microsatellite instability). Panel H is a consensus set of the groups in panels B-G that contained 34 cell lines in total with the overlap with DepMap shown [13]. Panel I shows groupings obtained using markers differentiating adenocarcinoma and mucinous adenocarcinoma portrayed as a spectrum of goblet-like properties [14]. Panel J shows the origin of the cell lines between the primary tumor (Prim) and metastatic sites (Met). ? Not clear. MSI+ cell lines. NS-no significant differences in *GPX2* expression levels among groups. Oversized dots indicate cell lines HT29 (red), SW480 (dark green) and SW620 (yellow). Stout arrow in panel G indicates that EMT is a strong property of the 1.1 class and the thin arrow that class 1.2 has the most inflammatory-like cell lines.

The key point is CRC tumors can be divided into up to six classes and in addition to differences in gene expression profiles, each category has a tendency for certain driver mutations, levels of copy number variation, CIMP profiles, MSS/MSI status and important for this discussion, difference in poor relapse free survival (RSF) and metastasis. The classification systems of the various groups while showing some discrepancies are similar enough that they were consolidated by Guinney et al into a consensus classification system (CMS; Fig 8H; involving groups in panels B-G) [49]. The availability of high throughput transcriptomics drove much of the classification, being relatable to cellular phenotypes and the clinical behavior of tumors. The consensus classification uses a neutral terminology (CMS1-4), however, this is sometime phrased as enterocyte- (Ent-), goblet- (Gob-, CMS3), transit amplifying- (TA-, CMS2), inflammatory- (Inf-, CMS1) and stem-like (Stem-, CMS4), suggesting gene expression level affinities to normal cell types of the colon (Figures 8A and F) [6,7]. CMS1 tumors tend to be hypermutated and is largely MSI and CIMP+, with *BRAF* mutations, and low copy number variation. CMS2 has high copy number variation, prevalent *APC* and *TP53* mutation and is MSS. CMS3 has a high prevalence of *KRAS* and *APC* mutations with some *TP53* mutation, moderate levels of CIMP+ and is MSS. CMS4 has high copy number variation, is MSS, and has moderate levels of mutation in *KRAS*, *APC* and *TP53*, and largely CIMP-. Guinney et al also provide an extensive analysis of differential properties for WNT/MYC target expression (CMS2), metabolic processes (CMS3, sugars, fatty acids), immune processes (CMS1), and EMT/TGF- β

activation (CMS4) among others. Consistent among the individual classification schemes and preserved in CMS is the finding of poor relapse free survival in CMS4 and, in Guinney et al, over all survival as well. CMS4 is also noted for metastasis, associated with high EMT pathway expression. CMS1 is associated with poor survival after relapse. These last factors will be important in the discussion of cell lines choices and it is worth double checking potential cell line choices for consistency with the MSS/MSI, Cimp+/-, copy number variation and mutation status properties of the related tumor class. All of this was intended to produce better prognosis criteria, precision therapies, and inform the process of cell line selection for pre-clinical testing of such therapies, Medico et al concentrating on the choice of cell lines (Fig 8A) [6]. An anomaly in the CRC nomenclature system of Medico et al (inherited from Sadanandam et al) is that in applying a system based on markers for normal cell types to cancer, some classical stem cell and WNT markers are associated with the TA-like class (*ASCL2*, *RNF43*, *ZNFR3*, *AXIN2*) (Figure 8A) [50]. Sadanandam et al stated that the main characteristics of the Stem-like CRC class were high expression of myoepithelial and mesenchymal genes and lack of differentiation. While they also say this CRC class has high WNT marker expression, this was not found to be a consistent property of the cell lines that were sorted into this category based on the stripped down iCMS marker set and was not supported by the findings of the CMS analysis for tumors of the CMS4 set (Figure 9) [15,49]. Lack of differentiation with high WNT marker status was oddly found in cell lines scattered across the Ent-, Gob-, Inf-, and TA-like classes.

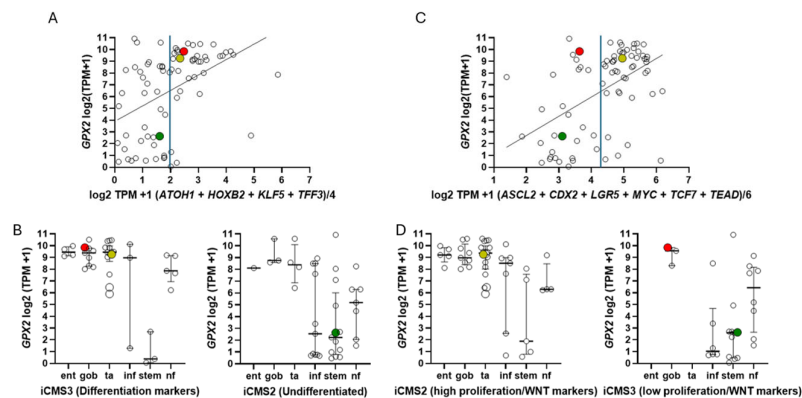


Figure 9. Panels A and B show the correlation of *GPX2* levels with the average of 4 differentiation markers (A) and a classification plot for cell lines with high or low expression of differentiation markers based in the iCMS system (B) [15]. In both panels HT29 is indicated in red, SW620 in yellow and SW480 in dark green. Panels C and D show the correlation of *GPX2* expression levels with the average of *ASCL2* + *CDX2* + *LGR5* + *MYC* + *TCF7* + *TEAD* expression in cell lines (C) and in a cell line classification plot with high or low proliferation/WNT marker cell lines (D). In panels A and C, the vertical blue lines denote the rather arbitrary division of low vs. high marker expression. While isogenic lines, SW480 and SW620, are separated in Panels B and D, the isogenic pair, GP2D and GP5D (large white circles), remain associated.

One notable outcome of the collective efforts was the relatively consistent classification of some cell lines into epithelial to mesenchymal transition (EMT)/TGFB pathway activation/undifferentiated/low WNT expression class (Figure 8). A characteristic of this set of cell lines in most classification schemes is low *GPX2* and *NOX1* expression (median levels- 4.35 TPM and 0.1 TPM vs. 588 TPM and 9.4 TPM in other classes, respectively). There is some inconsistency among the classification schemes for placement of cell lines, and low expressing lines are sometimes split among 2 or 3 classes within schemes (Figure 8A-I). For CMS, only 34 cell lines have been independently classified with 23 found in DepMap (Fig 8H). The results largely recapitulate Medico et al, so the

results of the Medico cell line classifications will be used for further discussion. Since the main goal was to observe how high and low *GPX2* expressing cell lines sorted based on the various CRC typing methods, extensive statistical analyses were not performed. For the Medico et al analysis, Ent-, Gob- and TA-like *GPX2* expression levels are not different and Inf- and Stem-like are not different. Inf and Stem are significantly different from Ent-, Gob- and TA-like (Figure 8A) ($p \leq 0.021$; Mann-Whitney; log2 transformed data directly from DepMap; PRISM 9.3). For CMS sets, CMS1 is not different from any other set; this classification corresponds to the Inf-like cell lines of Medico et al. [13,49]. CMS2 (largely TA-like) and CMS3 (mix of Gob-, TA-, and Inf-like) are not different, while both are different from CMS4, corresponding to a mix of the Stem-like and Inf-like sets of the Medico et al. classification (Figure 8H) ($p \leq 0.0056$). The mucinous adenocarcinoma marker analysis showed the least goblet set to be different from the core goblet and hybrid sets, which were not different (Figure 8I) ($p < 0.0002$). In the case of cell line origin, primary tumor or metastases, and MSI, no significant differences were found (Figure 8J).

Possible evidence of deep secretory-like derived cell lines (colon Paneth cell equivalent; possible candidates for high *GPX2* expression) in the DepMap set can be found using the differential markers indicated in Sasaki et al [33]. It is not clear whether such cells would have any tumorigenic capacity or whether any of the original differential markers of the postulated cell type would remain detectable when or if a transition like that of Paneth cells acquiring stem cell properties occurred. Presumably having acquired high WNT marker status while retaining some expression of differentiation markers, *CD24*, *REG4*, *MMP7*, *DLL4*, *EGF*, *SPDEF*, and *ATOH1*, they could be represented by 20 of 40 cell lines of the Medico et al classification scheme that overlap with DepMap, excluding stem-like and the NF categories, with *GPX2* levels >190 TPM and 1 stem-like cell line at 1 TPM (results not shown). This is pure speculation, as mice were studied for evidence of deep secretory cells, and they are not yet documented in human samples.

3.6. The Issue of Low *GPX2* Expressing CRC-Derived Cell Lines and Circumvention

Where do low expressing cell lines come from? In cancers of the GI tract *GPX2* expression is confined to uniform high levels, more so for COAD and READ (Figure 1A). In other cancers the range of expression is enormous, possibly explaining the establishment of some cell lines with low levels. Some low expressing cell lines from CRC have mutant APC and others with wild-type APC have high expression of at least one WNT ligand (DepMap) (Figure 4). Thus, the WNT dependence of *GPX2* expression, demonstrable in some CRC-derived cell lines and mice, does not seem to be the reason for low expression [51]. Analysis for *BRAF*, *KRAS*, *PIK3CA*, and *TP53* mutation did not reveal any association with *GPX2* expression levels in CRC-derived cell lines (DepMap; results not shown). *SMAD4* mutations are more frequent among the high *GPX2* expressing lines, but also found in intermediate and low expressing lines (11 of 14 cell lines with mutations, *GPX2* TPM > 64 ; 22% of high expressing lines, 11% of low expressing lines; $p=0.36$). *GPX2* expression levels are linked to NRF2 activation [2]. There aren't potent activating agents in standard culture media and this could be one factor in low *GPX2* expression. TP63 is one of the major drivers of *GPX2* expression in other tissues and cancers and there is sometimes a demonstrable correlation between TP63 and *GPX2* levels among cell lines derived from these other cancers (Figure 10) [52]. TP63 expression is notably absent in all but a few CRC-derived cell lines being replaced by WNT (Figure 10) [51].

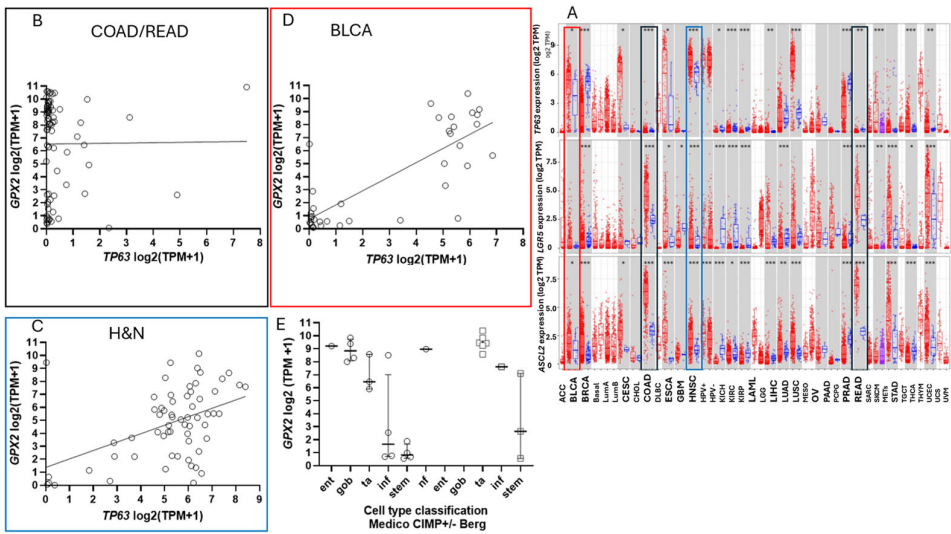


Figure 10. Normal and cancer tissues expression of *TP63* vs. *ASCL2* and *LGR5* by source. Black rectangles denote COAD and READ, red-BCLA and blue-H&N cancer (Panel A). Panels B shows the lack of *TP63* expression in CRC-derived lines, comparable to results in panel A, while panels C and D show generally higher expression of *TP63* in cell lines from H&N cancers and BLCA with some indication of a positive correlation with *GPX2* expression. E. Distribution of CIMP+ (circles) and CIMP- lines (squares) in Medico et al classes with *GPX2* levels.

Because the classification of cell lines was done after the fact, there will be discrepancies with the properties of the related tumor types. One of the more glaring issues is the excess of CIMP+ cell lines in the stem-like (CMS4) class. There is some trend for CIMP+ cell lines of the Inf- and Stem-like categories to have low *GPX2* expression (<6 TPM; 3 of 4 Inf-, and 4 of 5 Stem-like; note, CIMP+ is not a general property of the corresponding CMS4 class for Stem-like tumors and is a property of the corresponding CMS1 class for Inf-like tumors), while CIMP+ lines in the Ent-, Gob- and TA-like sets retain higher expression (≥60 TPM; Figure 10E). This might indicate suppression of *GPX2* gene expression by this repressive mark [53,54]. Two CIMP-, Stem-like lines also have low *GPX2* expression, so it is not clear what the relations are; CIMP+ in the Inf- and Stem-like categories specifically leading to low *GPX2* expression or more likely the condition of being Inf- or Stem-like leading to low *GPX2* expression as the primary effect while the CIMP+ condition is an incidental feature over-represented in CRC cell lines relative to tumors [13,53,54]. One criticism of cell lines is that there is a tendency for excess CpG methylation (a little less extreme in CRC-derived lines, 5-fold), although methylation sites found in CRC tend to be preserved. This was attributed to myriad effects of cell line derivation and cell culture conditions, to which loss of variability of CRC-derived lines compared to CRC was also blamed [53–55].

Based on the consistent EMT expression profile, it might be conjectured that the low *GPX2* expressing cell lines derive largely from metastases. This is not the case (Figure 8J). In fact, one study attempting to resolve the roles of *GPX2* in tumorigenesis and metastasis showed that while HT29 cells with *GPX2* knocked down showed poor growth as xenografts in mice, they could metastasize, although demonstrating a lower potential (*GPX2* levels in HT29 at 1010 TPM -averaging DepMap and THPA values; about 20-fold knockdown; *NOX1* 16 TPM) [56]. However, follow-up analysis showed that the metastatic tumors were a subset that either didn't fully respond to the knockdown or re-established higher *GPX2* expression in the process (IHC). Since HT29 cells were classified as CCS3 (Figure 8D, EMT set; poor relapse-free survival (RFS)), the authors went on to examine RFS in CCS3 cancers stratified by low and high *GPX2* expression, with high expression being unfavorable; questionable significance (TCGA/THPA does not confirm any prognostic relationship between *GPX2* and overall survival (OS), lumping all classes together; Figure 11).

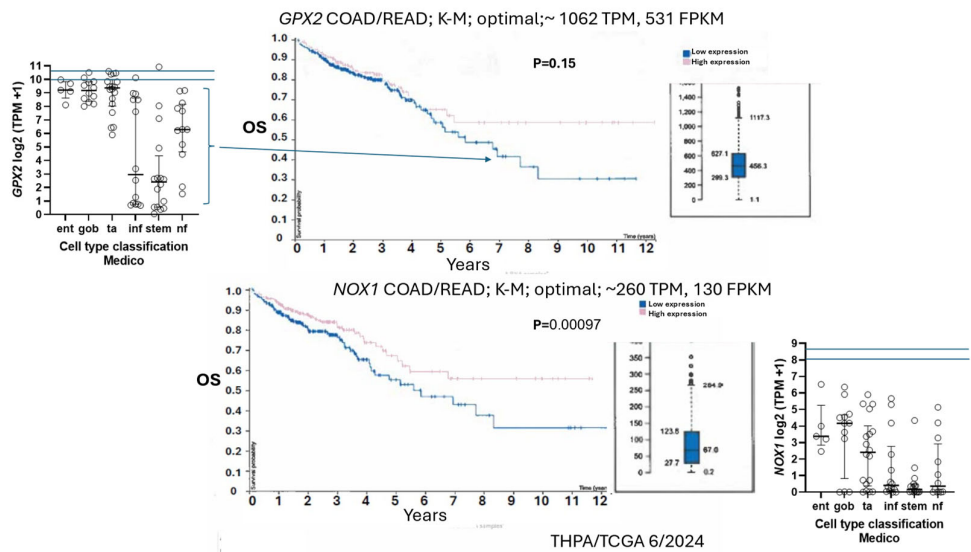


Figure 11. Kaplan-Meier plots from THPA (6/24) assessing OS prognosis based on *GPX2* and *NOX1* expression levels. The approximate cutoff levels are indicated as horizontal lines on the cell lines classification/expression levels plots. The upper line in the *GPX2* and *NOX1* plots are a correction to the TPM for the median tumor purity for CRC samples, made by multiplying the TPM (derived from TPM \approx FPKM \times 2, empirically determined conversion factor from ref. [1], by 1.6, correcting to a tumor purity of 100%, based on median tumor purity for CRC samples of 0.62 for comparison of tumor samples to cell line TPM values [1,44]. The bracketed cell lines in the *GPX2* plot are those with expression levels below the cutoff, although cell lines with values less than log₂ 8 are not really indicated by tumor data. The inserts are the distribution of tumor sample expression levels in FPKM.

They noted that the median level of *GPX2* expression in CCS3 CRC subsets was ~512 TPM (correcting for median tumor purity in CRC-820 TPM) much higher than the 5 TPM characteristic of most CCS3-type cell lines (CCS1 vs CCS2 and CCS3, $p \leq 0.0006$; Figure 8D) and that the variation in expression levels was only 4-5-fold among the CCS3 tumors [5]. The relevance of a 20-fold knockdown is questionable in the context of probing RFS. It might be reasonable as a therapeutic goal given the general negative impact on tumor growth and metastasis. HT29 is generally not classed into the EMT/TGFB pathway activation/ undifferentiated/low WNT expression set among classification schemes (Figure 8). The cell line is somewhat differentiated (some *MUC2*, *TFF3* expression and high *LYZ*; *ATOH1*, *HOXB2*, *KLF5*, *TFF3*; iCMS), has somewhat low TGFB pathway marker expression (*CCN1*, *ITGAV*, *TGFB*) and low EMT marker levels (*FN1*, *CDH2*, *VIM*) while sharing low WNT marker expression (*ASCL2*, *CDX2*, *LGR5*, *MYC*, *TCF7*, *TEAD*; iCMS) characteristic of the EMT sets (Figure 9) [57–59]. HT29 can undergo differentiation in culture with the substitution of galactose for glucose in the media, with some cells exhibiting goblet cell-like morphology and mucin production [60]. Slightly lower GPX activity in HT29 was associated with differentiation by 5-Fluorouracil [61]. This work predates wide-spread recognition of GPX isoenzymes so GPX1 and GPX2 were not separately examined. As a side note, galactose increases mitochondrial H₂O₂ release, hinting at a role for ROS in the differentiation process along with the reported lowering of GPX activity [62].

Nine cell lines with higher *GPX2* levels are classified by 2 or more of the schemes referenced in Figure 8 into the EMT/undifferentiated categories; CACO2 (3 schemes), CL40 (2 schemes), LOVO (3 schemes), LS180 (2 schemes), SKCO1 (2 schemes), SNU1040 (3 schemes), SW620 (5 schemes), SNU503 (5 schemes), and SNU1033 (3 schemes) with *GPX2* expression levels of 118, 918, 290, 296, 1711, 1020, 648, 2040, and 318 TPM, respectively (average, DepMap and THPA) (Table 1) [7–12].

Table 1. Color coding indicates cell lines appearing in multiple classifications; see Figure 8 and main text. MSS-microsatellite stable; MET-cell line derived from metastasis.

TABLE 1 EMT/undifferentiated cell lines with high GPX2 expression								
Sadanandam	Marisa	Budinska	de Saúsa E Melo	Roepman	Schlicker	MSS	MET	APC mut
SNU503	T84	SNU1040	SNU1040	SNU1040	CACO2	SNU1033	SNU503	SNU1033
SW620	SNU407	SNO503	HT29	CL40	LOVO	SNU503	LOVO	SNU503
	LOVO	CACO2	SNU81	SKCO1	SW1417	CACO2	SW620	LOVO
	LS180	LOVO	SNU1033	SNU1033	SW620	SW620	SKCO1	SW620
	HT115	C84	SNU503	SNU503		CL40		SKCO1
	LS411N	SW620	SW620	LS180		SKCO1		CL40
	SNUC41			SW620				SNU1040
	SNU1033			COLO201				
	SNU503							
	CACO2							
	CL40							
	SNU61							
	SNU1460							
	SKCO1							
	SNUC4							
	CL34							

There are discrepancies with and between Medico et al/Sadanandam et al; LS180 and SKCO1 are classified as Gob- by both, while SNU1040 is split Ent-/Gob- and SW620 is split TA-/Stem-like. This appraisal might be further limited by the suggestion that MSS tumors of the CMS4 class are most likely to metastasize, leaving CACO2, CL40, SKCO1 (metastasis), SNU503 (metastasis), SNU1033 and SW620 (metastasis) as representative cell lines with potentially high metastatic potential [15]. However, LOVO (MSI; metastasis) and HCT116 (MSI; low *GPX2* expression; identified in the EMT/undifferentiated classes in all schemes) were found to have metastatic potential in a mouse xenograft model rivaling that of SW620, a cell line derived from a metastasis with metastatic potential demonstrable in 2 mouse model studies [63,64]. There may be a few seemingly EMT/TGFB pathway activation/undifferentiated/low WNT expression category cell lines with *GPX2* expression levels approaching tumor levels, which seems desirable in testing the role of *GPX2* in this class of tumors based on the commonly used suppression of *GPX2* expression in cells lines as a mode of study; notably, CACO2, SNU503, SNU1033, and SW620 for Stem-like cell lines. CL40 and LOVO might be as appropriate as high expressing lines for the Inf-like class (Medico et al/Sadanandam et al.) [6,7]. Within these two classifications, paired high and low expressing lines may represent cases where the assumption that a major difference is in *GPX2* levels is somewhat justified and the use of forced over-expression and silencing might yield relevant findings, with the caveats that very low expression in cell lines is not yet supported by CRC expression levels and differences in expression levels of *CST6*, *EGFR*, *MYC*, *TP53*, *VIM* and other putative redox sensitive proteins may be present.

3.7. How Different are CRC-Derived Cell Lines Sorted into Classes by Medico et al for Expression Levels of Genes Encoding Proteins with Reactive Sulfhydryl Groups?

While the CRC-derived cell line sorting process is largely based on differential gene expression profiles, just how different are the cell lines for mRNA levels of proteins with known or suspected reactive sulfhydryl groups that impact function when oxidized? While the list of potentially affected proteins is large and growing, a few key proteins were selected for evaluation in DepMAP [2,65]. Uniformity and variability in expression levels were found across the groups of cell lines in the Medico et al classification scheme (DepMap) [66–68]. 14-3-3-gamma (*YWHA*G), c-ABL (*ABL*1), *ASK*1, *AKT*1, *AKT*2, *HIF*1A, *JNK*2, *JUN*2, *ERK*2, *KEAP*1, MAPKs, *NFKB*, Nucleoredoxin (*NXN*), p38 (*MAPK*14), protein kinase A (*PRKACA*), *PTEN*, *PTPN*1B, *SHP*1, *SHP*2, *SRC*, and *STAT*3 have uniform expression levels across classes. Cystatin B (*CST*6), *EGFR*, *MYC*, *TP53*, and vimentin (*VIM*) show skewing for certain classifications; *EGFR* levels tend to be slightly higher (2-4-fold) in low *GPX2*

expressing lines, enriched in Stem- and Inf-like cell lines, some with high *GPX2* levels, although a few very low expressing lines have nearly zero levels; 4-fold lower expression of *TP53* is associated with the high *GPX2* expressing TA-like cell lines compared to the other classes; high levels of *CST6*, 16-fold or more, are generally associated with the low *GPX2* expressing Stem- and Inf-like classes; *VIM* is slightly skewed in favor of low *GPX2* expressing cell lines (2-4-fold), with 5 Stem-like low expressing lines and SW620, a high expressing Stem-like line (most schemes; TA-like in Medico et al; Figure 8 and Table 1), showing roughly 8-32-fold higher expression than the bulk. *MYC* shows slight skewing (2-fold) in favor of low *GPX2* expressing Stem- and Inf-like cell lines. It could be argued that for most redox sensitive proteins upstream of major pathways there would be a uniform impact of ROOH modulation across CRC classifications with a few exceptions and, that in many cases mRNA levels do not reflect functionality of the proteins. However, this exercise can only be efficiently managed using the available mRNA expression data. *CST6*, *EGFR* and *VIM* might be exceptions, where mRNA levels reflect function. The other point is that the downstream protein profiles may be variable among the CRC-derived cell lines based on the classification method; Medico et al used 783 differentially expressed genes in their profiling [6]. That is, *GPX2* expression level differences generally occur in a context of broad gene level expression differences and not in isolation. Forced overexpression and silencing of *GPX2* in one or a few cell lines may not reflect this situation, biasing results in favor of the impact of the consensus redox sensitive proteins and downstream pathways, presuming these remain largely unaffected by the manipulations. However, it was shown above that there may be a few suitable pairings of high and low *GPX2* expressing lines for the Stem- and Inf-like classes. While the argument has been made that *GPX1*, *PRDXs* and *CAT* impact should also be accounted for, a broader systems approach may be required to understand how *GPX2* expression alterations (and the other antioxidants) fit into the scheme of CRC and what drives the large-scale gene expression changes that include *GPX2*.

3.8. Intratumor Variation, Plasticity of Cell Components, Stromal Cell Interactions and *GPX2*

There is a general trend for high *GPX2* expression in normal colorectal samples at the tissue level (Figure 2; ?). This appears to be achieved with a limited zone of high expression at the base of the glands [29–31]. There is disagreement on whether a low level of *GPX2* expression extends to the top of the gland. While one study mentioned above showed individual CRC tumor cells to have elevated *GPX2* protein levels relative to expressing normal cells in a somewhat small set, it is not clear that this is universal in CRC samples [29]. THPA *GPX2* IHC sets show a few tumors where *GPX2* intensity is quite variable in the epithelium, with portions of tumors showing almost no staining adjacent to strongly stained sections and other lesser variation is also observed (Figure 7A and B). The unstained or lightly stained portions are likely the normal epithelial cell component of tumors (not considered as a source of CRC-derived cell lines) that occurs at an average 24% of epithelial cells isolated from tumors in the iCMS analysis, presumably having low levels of *GPX2* relative to the tumor cells [15]. The sample shown in Figure 7B might represent a tumor with a normal component above this average, while panels C and D seem to be below the average. The reverse staining pattern is unlikely given the low numbers of normal high *GPX2* expressing cells, unless alterations in the tumor stromal cellular environment promote aberrant growth of normal cells along the lines of Paneth cells in tissue injury (postulated deep secretory cells in colon) [30,32–35]. Whether any of this variation relates to the tumor cell component has not been systematically studied so we would have no idea about how prevalent it is as a pattern and whether it relates to CRC subtypes [15]. If this involves the tumor component, a conjecture is the CMS4/CCS3 class may be more represented explaining the propensity for low *GPX2* expressing cell lines from this class (Figure 8). This can be readily tested, and if the variation involves the tumor cell component this could justify the use of low expressing cell lines in studies.

GPX2 expression in primary tumors may be dependent on factors secreted by CAFs and other stromal cells and portions of the tumors may lack sufficient contact with these cells to support *GPX2* expression. Growth in primary sites seems to be somewhat dependent on stromal support [69–71]. Tumor cells seem to be plastic, altering their properties, which might include cellular factors affecting

GPX2 expression, separate from stromal factors [38]. Stromal dependency, while possible, is countered by the more general tendency of cell lines to retain *GPX2* expression in standard culture conditions and the idea that metastases can be established in the absence of stromal support (DepMap) [38]. It is estimated that only 10-15% of CRCs can supply cells that will establish cultures [72]. A general requirement for stromal support in primary tumors and re-establishment of this after metastases have seeded may explain part of this failure. In standard isolation of CRC-derived and other cell lines, fibroblasts are the main contaminant, and it takes several passages to eliminate them, or they may be employed as feeders until overgrowth is evident as a sign of independence of the CRC epithelial cells [73]. It is possible that a transition from stromal dependency to independent growth occurs by adaptation or selection as the fibroblasts are eliminated [73]. There are no broadly consistent patterns of gene expression among cell lines as evidenced by the stratification into as many as 6 subtypes based on differences in expression, driver mutations and epigenetic modulation patterns (Figure 8). It is possible to find many different combinations of driver and passenger mutations in the list of 77 CRC cell lines in DepMap [47]. Thus, the basis for culturing success is elusive and clearly has nothing to do with *GPX2* expression either directly or in the context of cell properties favoring *GPX2* expression.

A side project on *GPX2* expression could include adding CRC stromal factors back to cultures, or better, using stromal cells as feeders of the low expressing lines looking for induction of *GPX2* levels to the range of 100-2000 TPM [73]. There is evidence that high *TGFB* expression in tumor cells, somewhat greater in low *GPX2* expressing cell lines (DepMap), exerts a paracrine effect on the surrounding stromal cells, which in turn enhances metastasis of fibrotic-type CRC tumors [74]. This would have to be linked to a broader examination of gene expression alterations to avoid overinterpretation for any impact of *GPX2* on subsequent cell line properties. Another point raised by the substitution of galactose for glucose in the culture of HT29 and the change in properties is the presence of additional carbon sources for intestinal cells, short-chain fatty acids and glutamate [75,76]. Addition of short-chained fatty acids to culture media seems to evoke changes in cell line properties with a generally negative impact on growth [77]. Retinoic acid was found to increase *GPX2* expression in cell lines, in particular, MCF7, a line that barely expresses, *GPX2* under basal culture conditions [2].

As a general statement, based on Figure 2A and the IHC in Figure 7 (several other similar examples can be found at THPA), there is an expectation that CRC tumors will have high *GPX2* expression both at the tissue level and throughout the epithelium. A study attempting to understand changes in epithelial properties as CRC tumors develop from polyps found that Stem-like cells showed a consistent increase in *GPX2* levels by 3-5-fold as the malignancy progressed, relative to normal Stem-like cells, with elevation evident in the polyps [32]. This change in expression levels is coincident with the 3-5-fold increase in *GPX2* protein levels expression found by Brzozowa-Zasada et al in the comparison of high expressing normal cells and tumors [29]. *ASCL2* and *OLFM4*, stem cell markers, show increases in levels in the same cell population while *HNF4A* has a bi-phasic pattern, decreasing from normal to the polyp phase then increasing dramatically in the CRC stage. *HNF4A* shows a slightly stronger correlation (R-squared, 0.48) with *GPX2* levels in cell lines than *ASCL2* (R-squared, 0.31) and *OLFM4* (R-squared, 0.18), but for all the residual deviation from the regression line is huge for a large fraction of lines. *GPX2* has been occasionally proposed as a stem cell marker, so that correlations to other such markers in CRC samples (*AXIN2*, *RNF43*, *ZNRF3*, and *SOX9*; no correlation with cell line *GPX2*, although all are up regulated in CRC) have been made before and associations with WNT proposed and established [32,50,51]. For *HNF4A*, the pattern is explained as yielding to WNT dependency early on, it is an inhibitor of WNT, and fueling CRC later [32,78]. The number of independent samples in this study was very small, only three CRC samples, and unable to cover the spectrum of CRC classes. However, there is no indication that *GPX2* levels should be low in tumor samples.

Pathways that drive high *GPX2* expression, indicated by results from tumors, could be replicated in high expressing cell lines with the provision that there may be some dependence on stromal factors in tumor expression that would not be carried over to cell culture. There are many

pathways that impact *GPX2* levels, so finding a really strong correlation with any single gene or pathway is unlikely [2]. Esworthy, Doroshow and Chu presented an overview of genes and pathways that influence *GPX2* levels and the count was at least 14, at that time [2]. *NFR2*, *WNT*, *TP63*, and *RARE* have been mentioned above. Others, including *CD13*, *ETS2*, *FOXO1*, *FOXM1*, *OXR1*, *PI3K/AKT* and *STAT3* do not show a strong correlation with *GPX2* expression in cell lines and *NKX3-1* is not expressed in colon. *FOXM1* is up regulated in CRC and *FOXO1*, *EST1* and particularly, *CD13* (*ANPEP*), are down regulated (TIMER2.0). *HIPPO/YAP* is unique, reported to strongly suppress *GPX2* expression. Like *HNF4A*, which is in a cooperative feed-back loop of expression with *HNF1A* and therefore involved with the *NOTCH* pathway, a few master regulators of CRC pathways, *MYB*, *REG4*, and *HIPPO/YAP* show a reasonably strong correlation with *GPX2* expression, markers of *YAP* activation and related genes, a negative effect [78,79]. *MYB*, which is expressed in other tissues, shows marginal up regulation in CRC with a fairly strong positive correlation with *GPX2* in cell lines (R-squared, 0.498) [80]. *REG4* is expressed in normal GI-tract tissues and is not uniformly up regulated in CRC [33]. The correlation with *GPX2* in cell lines is weaker (R-squared, 0.22). *YAP1* is widely expressed at uniform levels and is not up regulated in CRC. Expression levels of *YAP1* and *TAZ* (co-activator) are also uniform across cell lines classes. Markers of *YAP/TAZ* pathway activation show various degrees of negative correlation with *GPX2* expression in cell lines, *CCN1* (*CYR61*) with an R-squared of 0.64 (DepMap). *CCN1* is also a marker of the *TGF β* pathway, providing two possible reasons for the strong correlation [17]. *YAP* activation might occur in cell culture, as it senses mechanical cues related to cell shape and cell spreading [13,81]. Thus, low *GPX2* expression in the Stem-like category of cell lines could reflect an impact of conventional cell culture. *YAP* activation was tied to down regulation of *GPX2* in LUSC through *TP63*, which is not expressed in colon or CRC-derived lines, requiring another mechanism of action (Figure 10)[82,83].

3.9. *NOX1* in CRC and Links to *GPX2*

Expression of *NOX1* among CRC-derived cell lines shows a similar pattern of variable expression with a greater proportion of lines exhibiting nearly zero expression relative to *GPX2* (Figure 5). Very high *NOX1* levels are a unique property of normal colon/rectum and CRC with a narrow range of variability in expression levels in normal and a very broad range of expression in CRC, so that low expression in cell lines is not discrepant as with *GPX2* (Figures 2 and 5). *NOXO1*, but not *NOXA1*, shares unique high expression in colon/rectum (TIMER2.0). The only linkage in co-expression might be that high *NOX1* levels are associated with high *GPX2* levels as observed in cell lines (DepMap) (Figure 5). *NOX1* mRNA and *NOX1* protein expression occurs at crypt/gland base like *GPX2* [26–28]. Whether co-expression occurs in the same normal cells is not clear. The absence of *NOX1* (*NOX1*-KO) in mouse colon has some impact on the distribution of differentiated cells (goblet cells) and proliferating cells, favoring differentiation [84]. This suggests that *NOX1* might be expressed in the TA compartment, which is the cell line assignment of Medico et al [6]. In the absence of *GPX1* and *GPX2*, *NOX1* activity produces pathology in the form of excess crypt/gland apoptosis, which may be dependent on the presence of microbiota and the composition, providing one rationale for very high *GPX2* in colon/rectum [24,25,85,86]. The pattern in CRC-derived cell lines shows co-expression of *GPX2* and *NOX1* is likely in cancer as property of several, but not all, high *WNT* expressing and differentiated cell lines (Figures 5 and 11). Unlike *GPX2*, *NOX1* expression levels are considered prognostic by TCGA/THPA for OS, as of 6/2024 (Figure 11). *NOX1* stratification yields a significant effect on OS with high expression being favorable. The cut-off for the prognosis call is ~262 TPM (420 TPM, correcting for median purity values), which is greater than 2-times more than even the highest expressing cell line (Figure 11). There is the question of whether *NOX1* would be active in any cell lines. One paper tried to address this by showing that a handful of CRC-derived cell lines at the top of the expression range might be capable of significant spontaneous and PMA-induced production of superoxide from *NOX1* [87]. The knockdown of antioxidant enzyme expression in CRC cell lines might reveal activity as increased levels of apoptosis that should relate to *NOX1* levels, if true.

3.10. Low GPX2 Expression in Cell Lines, Redux

Returning to the issue of how to select cell lines for pre-clinical studies of the involvement of GPX2 in CRC, a major question is whether to use low expressing lines for studies. The observation of low GPX2 protein levels in portions of tumor samples may be insignificant as a consideration in CRC studies, representing normal cells, and low expressing cell lines may be of little interest and utility in CRC studies despite their otherwise interesting properties. Plasticity is noted to occur in cancer stem cells. Single-cell analysis found a low incidence of multiple subtype signatures in individual tumors, probably based on one sample per tumor [15]. Another study showed that multiple sampling within metastatic tumors revealed unrecognized heterogeneity [88]. Plasticity in CRC tumor stem cells, the finding of a few tumors with multiple signatures and the likelihood of regional variation within tumors suggests that the coexistence of low and high expressing cells in tumors could occur with selection for alternate states a possibility [38]. GPX2 expression might not be an essential feature of tumor progress, only a legacy of expression potential. This might thwart efforts to use modulation of GPX2 levels to achieve therapeutic ends. There is possibly only one example of such behavior in established cell lines and there is a trend for consistent increases in GPX2 expression along what is termed the malignancy continuum suggesting a singular upward arc during tumorigenesis for Stem- and TA-like cell types [32]. Cell lines, SW480 and SW620 are isogenic, the former isolated from the primary tumor and the latter from a lymph node metastasis [89]. Among the many differences, the GPX2 level in SW480 is ~5.7 TPM and in SW620 is ~648 TPM (averaging results from DepMap and THPA), SW620 has much higher expression of *LGR5* and *ASCL2*, and slightly greater signs of differentiation, while they share very low *NOX1* expression (Figures 5 and 9). In two studies of metastatic potential in mice, SW620 demonstrated a much greater ability than SW480 [63,64]. Comparing GPX2 levels (DepMap) to the reported metastatic ability among 6 cell lines used in one of the two studies did not show any trend, suggesting other properties of those lines and possibly SW480 and SW620 were responsible [64]. The paired lines, GP2D and GP5D, retain similar, somewhat low GPX2 levels, 87 TPM (83 TPM averaging DepMap and THPA) and 59 TPM, similar proliferation and differentiation marker status and low *NOX1* levels (Figures 5, 9B and 9D). Both were derived from the same primary tumor and exhibit some different properties in culture, including responses to EGF ligands and spontaneous vs. induced EMT [90]. Supplementation of EGF ligands to culture media could be one more item to explore for impact on GPX2 expression in addition to short-chained fatty acids and retinoic acid. Other pairings exist or may exist; however, one member of the pair doesn't appear in the DepMap or THPA databases or the suggested pairing is based on later supposition and not the original description (DLD1, HCT15; primary; similar low GPX2 and *NOX1* expression; multiple sources support pairing) (Figure 5) [6]. CRC organoids might provide a means to explore the plasticity issue [56].

4. Summary

In general, it appears that cell line selection for studies should account for CRC subtypes, selecting across broad gene expression profiles, different driver mutations and epigenetic modulation. There is sufficient information and enough cell lines typed to account for most of the variation across CRC classes (DepMap, THPA) [6–15,47]. Medico et al sorted 150 cell lines and investigators could independently measure relative GPX2 expression in the lines not covered by DepMap to include other combinations. Note that RT-PCR results from individual studies are often at variance with DepMap/THPA [1]. An investigation into this suggested that DepMap/THPA values were generally reliable guides for predicting GPX protein and activity levels in cell lines at a gross level, while the differences observed between DepMap and THPA reveal uncertainty for fine distinctions (Figure 4C and E) [1]. It is recommended to use at least two cell lines, one low and one high expressing, from the DepMap collection to provide reference values when looking at other cell lines [1]. A study might need to use four or more lines to sufficiently account for the major variation among CRC classes, depending on the goals.

Classification systems tend to place high GPX2 expressing and low GPX2 expressing cell lines into different categories, and this may restrict the implementation of the common research scheme of

overexpression in low expressing lines and knockdown of high expressing lines (Table 1), although some methods place both high and low expressing lines across classifications (Figure 8). As to using cell lines to model cancer etiology, the top-down model would suggest a likelihood that the source cells lacked *GPX2* expression and gained it as a function of progress along the malignancy continuum [32,39]. The bottom-up model does not guarantee that the source cells had high expression, however, this is a possibility and cells in this area might be likely to acquire *GPX2* expression in an early hyperplasia phase based on evidence that NOX1 up-regulation and activation is involved in the shift of gland base stem cells from quiescence to active cycling [91,92]. As of this time, there is no clear association between the few high expressing cells found at the gland base and CRC. Until the plasticity and stromal contribution issues are resolved, cell lines with very low *GPX2* expression are something of an enigma, particularly with the trend for fully developed CRC to have higher *GPX2* expression than its antecedents [32]. Cells shed from primary tumors may lack *LGR5* expression and regain it after they seed metastases [38]. Likewise, it seems possible this pattern might occur for *GPX2* based on the study of Emmink et al [56]. The example of SW480 and SW620 suggests that primary tumors could have low expressing cells and still produce high expressing metastases [63,64,83]. This might reflect heterogeneity in the primary, plasticity and the malignancy continuum arc as it relates to metastasis or just be an artifact of cell culture [55]. There is no uniform difference in *GPX2* expression patterns for cell lines isolated from primary and metastatic sites (Figure 8J). In lung adenocarcinoma, lower RFS seems to be related to persister cells that maintain higher levels of *GPX2* expression, which may be essential for the resilience of this cell type both in the quiescent and cycling states [93,94]. This is consistent with the findings that higher *GPX2* expression is associated with worse RFS in CCS3 CRC and inconsistent with the apparent isolation of low expressing cell lines from CCS3 tumors (Figure 8) [56].

Based on the model for GPX1/2 action relative to the postulated PRDX relay (Figure 1), the main factors in CRC are ROOH signaling in which NOX1 may have a major role, the GPX: PRDX ratio that would modulate the ROOH signaling potential and the protein expression pattern of the cell type that would dictate the response to the ROOH signal. Another consideration is that PRDX reduction may be hampered by slow interaction with relay targets and a possible bottleneck with thioredoxin/thioredoxin reductase resulting in temporary inactivation [4,95–97]. GPXs may fill in for the loss of antioxidant capacity while PRDXs regenerate.

The relative TPM levels among GPXs and PRDXs are intended only as rough guides to comparative function, a starting point for thinking about where *GPX2* might have a significant role and for choosing cell lines. PRDX1 and PRDX2 are estimated to have initial ROOH rate constants 2-5-fold over that of GPXs, while PRDX5 is estimated to be 10-100-fold less efficient [3,4]. This is coupled with the generally greater abundance of PRDXs [1]. The TPM counts can be modified based on these considerations to more realistically model relative antioxidant potency. On the other hand, the estimates of relative normal colon/rectum and COAD/READ tissue levels underrepresent *GPX2* (20-27% total TPM) by including expression of *PRDX1-6* and *GPX1* outside of the epithelial compartments; *GPX1* and *PRDX1-6*, follow the pattern of β -actin in Figure 6. Attempting to correct for this revises the estimates for *GPX2* to 35-50% of total TPM; downplaying the contribution of PRDX5 by 10-fold and allowing the higher rate constants for PRDX1 and PRDX2, suggests *GPX2* could be nearly at parity with PRDXs for ROOH consumption. In much of the normal epithelium high *GPX2* expression is not found and in portions of some tumors this is seen as well, yielding additional upward revision (Figures 1B, 3 and 7) [1,29,30]. *GPX2* levels equal to or greater than 35% of total antioxidant gene TPM are found in 10% of CRC-derived cell lines (up to 45%) (Figure 4). Applying the rate constant corrections to 5 of the top expressing cell lines suggests that *GPX2* could account for 20-27% of total ROOH reduction, up to 30% including *GPX1*. The fraction of ROOH going through GPX1/2 is significant because it could divert ROOH from signaling via the PRDX relay (Figure 1A). Alteration of *GPX2* levels is significant if it actually impacts the fraction of the total ROOH reduction capacity; that is no commensurate changes occur among PRDXs. This is why information on GPX1/2 and PRDX1-6 is required in these studies to assess the actual impact. In the

case of NOX1, if the predominant signaling is based on superoxide and H₂O₂ is only a byproduct, having high levels of GPX1/2 could prevent unintended signaling (Figure 1B) [98].

Factoring in *NOX1*, possibly no CRC-derived cell lines would qualify for selection in studies (Figure 11). These considerations add more levels of uncertainty in evaluating the possible role of GPX2 in cancer and general antioxidant function, particularly using cell lines, which by and large fail to replicate the high *GPX2* and *NOX1* levels of COAD/READ. CRC and ESCA/STAD are among the better cases for finding representative cell lines [1]. LIHC-derived cell lines show vastly lower *GPX2* levels than the tumors (Figure 2).

5. Conclusions

Finally, no definitive answer can be provided on how to study *GPX2* in CRC. One question is whether forced overexpression or silencing of *GPX2* is the same as alterations to cells that result in changes to *GPX2* levels, where there might be changes to the potential target proteins as well. The work of Emmink, et al suggests that suppressing *GPX2* in cancers with high expression (>820 TPM; based on HT29; the levels in the CRC-derived organoids used were not provided with the default from databases being high expression) might suppress growth and metastatic potential [56]. The work involved both the HT29 cell line and organoids cultures from 2 CRC subjects, giving it some veracity. They did not provide any information on CRC subtypes of the source for the organoids and the in vitro culture conditions of the organoids were not typical, lacking WNT activators, Noggin, and RSPO, etc. [99]. There are several other *GPX2* and CRC papers that collectively suggest that *GPX2* levels should impact CRC progress. More generally in experimental work high expression seems to favor tumor progress with a range of proposed pathways being impacted [1]. There are some conflicting findings-high expression repressing the prostaglandin and lipoxygenase pathways and subsequent growth, in vitro, while the same condition supported tumor growth in mice and suppressed mutation which might disfavor immune check point therapies; low expression suppressed azoxymethane induced-CRC in mice [100,101]. Prognosis analyses conflict with TCGA/THPA suggesting no impact on OS, collectively, while focusing on the EMT subtype suggests high *GPX2* expression in this subtype may yield poor prognosis for RFS (Figure 11) [7–15,56]. CRC is complex and the issues of plasticity and stromal involvement further complicate the situation, suggesting that multiple pathways could be impacted by *GPX2* action and contribute to mixed research results. If it is worth the effort, research must account for the complexity and current uncertainties surrounding CRC and this would involve careful selection criteria for cell lines that would go beyond *GPX2* levels and employ the use of multiple lines in any single study. Alternatively, reading this report might convince many that using cell lines is not a good approach to understanding CRC or the role of *GPX2*.

Author Contributions: Conceptualization, Methodology, Formal Analysis, Investigation, Writing–Original Draft Preparation, Writing–Review & Editing, Visualization, RSE

Funding: none

Conflicts of interest: none

IRB: not required.

Data availability: All data can be found in the referenced public databases and cited papers; inquiries can be made to RSE; email, sesworthy@coh.org

Acknowledgements: The suggestions of Drs. Sarah P. Short, Anna P. Kipp and Fong-Fong Chu were incorporated into the text and the author thanks them for their valuable input.

References

1. Esworthy, RS, Chu, FF Using Information from Public Databases to Critically Evaluate Studies Linking the Antioxidant Enzyme Selenium-Dependent Glutathione Peroxidase 2 (GPX2) to Cancer. *BioMedInformatics*. 2023 Nov; 3: 985-1014 10.3390/biomedinformatics3040060.

2. Esworthy RS, Doroshow JH, Chu FF. The beginning of GPX2 and 30 years later. *Free Radic Biol Med*. 2022 Aug 1;188:419-433. doi: 10.1016/j.freeradbiomed.2022.06.232. Epub 2022 Jul 5. PMID: 35803440; PMCID: PMC9341242.
3. Averill-Bates D. Reactive oxygen species and cell signaling. Review. *Biochim Biophys Acta Mol Cell Res*. 2024 Feb;1871(2):119573. doi: 10.1016/j.bbamcr.2023.119573. Epub 2023 Nov 8. PMID: 37949302.
4. Villar SF, Corrales-González L, Márquez de Los Santos B, Dalla Rizza J, Zeida A, Denicola A, Ferrer-Sueta G. Kinetic and structural assessment of the reduction of human 2-Cys peroxiredoxins by thioredoxins. *FEBS J*. 2024 Feb;291(4):778-794. doi: 10.1111/febs.17006. Epub 2023 Dec 7. PMID: 37985387.
5. Xu H, Hu C, Wang Y, Shi Y, Yuan L, Xu J, Zhang Y, Chen J, Wei Q, Qin J, Xu Z, Cheng X. Glutathione peroxidase 2 knockdown suppresses gastric cancer progression and metastasis via regulation of kynurenine metabolism. *Oncogene*. 2023 Jun;42(24):1994-2006. doi: 10.1038/s41388-023-02708-4. Epub 2023 May 3. PMID: 37138031; PMCID: PMC10256612.
6. Medico E, Russo M, Picco G, Cancelliere C, Valtorta E, Corti G, Buscarino M, Isella C, Lamba S, Martinoglio B, Veronese S, Siena S, Sartore-Bianchi A, Beccuti M, Mottotese M, Linnebacher M, Cordero F, Di Nicolantonio F, Bardelli A. The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. *Nat Commun*. 2015 Apr 30;6:7002. doi: 10.1038/ncomms8002. PMID: 25926053.
7. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wulschleger S, Ostos LC, Lannon WA, Grotzinger C, Del Rio M, Lhermitte B, Olshen AB, Wiedenmann B, Cantley LC, Gray JW, Hanahan D. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med*. 2013 May;19(5):619-25. doi: 10.1038/nm.3175. Epub 2013 Apr 14. PMID: 23584089; PMCID: PMC3774607.
8. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi MC, Schiappa R, Guenot D, Ayadi M, Kirzin S, Chazal M, Fléjou JF, Benchimol D, Berger A, Lagarde A, Pencreach E, Piard F, Elias D, Parc Y, Olschwang S, Milano G, Laurent-Puig P, Boige V. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med*. 2013;10(5):e1001453. doi: 10.1371/journal.pmed.1001453. Epub 2013 May 21. PMID: 23700391; PMCID: PMC3660251.
9. Budinska E, Popovici V, Tejpar S, D'Ario G, Lapique N, Sikora KO, Di Narzo AF, Yan P, Hodgson JG, Weinrich S, Bosman F, Roth A, Delorenzi M. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J Pathol*. 2013 Sep;231(1):63-76. doi: 10.1002/path.4212. Epub 2013 Jul 8. PMID: 23836465; PMCID: PMC3840702.
10. De Sousa E Melo F, Wang X, Jansen M, Fessler E, Trinh A, de Rooij LP, de Jong JH, de Boer OJ, van Leersum R, Bijlsma MF, Rodermond H, van der Heijden M, van Noesel CJ, Tuynman JB, Dekker E, Markowitz F, Medema JP, Vermeulen L. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat Med*. 2013 May;19(5):614-8. doi: 10.1038/nm.3174. Epub 2013 Apr 14. PMID: 23584090.
11. Roepman P, Schlicker A, Tabernero J, Majewski I, Tian S, Moreno V, Snel MH, Chresta CM, Rosenberg R, Nitsche U, Macarulla T, Capella G, Salazar R, Orphanides G, Wessels LF, Bernards R, Simon IM. Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *Int J Cancer*. 2014 Feb 1;134(3):552-62. doi: 10.1002/ijc.28387. Epub 2013 Sep 13. PMID: 23852808; PMCID: PMC4234005.
12. Schlicker A, Beran G, Chresta CM, McWalter G, Pritchard A, Weston S, Runswick S, Davenport S, Heathcote K, Castro DA, Orphanides G, French T, Wessels LF. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. *BMC Med Genomics*. 2012 Dec 31;5:66. doi: 10.1186/1755-8794-5-66. PMID: 23272949; PMCID: PMC3543849.
13. Berg KCG, Eide PW, Eilertsen IA, Johannessen B, Bruun J, Danielsen SA, Bjørnslett M, Meza-Zepeda LA, Eknæs M, Lind GE, Myklebost O, Skotheim RI, Svein A, Lothe RA. Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. *Mol Cancer*. 2017 Jul 6;16(1):116. doi: 10.1186/s12943-017-0691-y. PMID: 28683746; PMCID: PMC5498998.
14. Hu FJ, Li YJ, Zhang L, Ji DB, Liu XZ, Chen YJ, Wang L, Wu AW. Single-cell profiling reveals differences between human classical adenocarcinoma and mucinous adenocarcinoma. *Commun Biol*. 2023 Jan 23;6(1):85. doi: 10.1038/s42003-023-04441-w. PMID: 36690709; PMCID: PMC9870908.
15. Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, Eng CLP, Macalinao DC, Kahraman M, Srinivasan H, Lakshmanan V, Verbandt S, Tsantoulis P, Gunn N, Venkatesh PN, Poh ZW, Nahar R, Oh HLJ, Loo JM, Chia S, Cheow LF, Cheruba E, Wong MT, Kua L, Chua C, Nguyen A, Golovan J, Gan A, Lim WJ, Guo YA, Yap CK, Tay B, Hong Y, Chong DQ, Chok AY, Park WY, Han S, Chang MH, Seow-En I, Fu C, Mathew R, Toh EL, Hong LZ, Skanderup AJ, DasGupta R, Ong CJ, Lim KH, Tan EKW, Koo SL, Leow WQ, Tejpar S, Prabhakar S, Tan IB. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat Genet*. 2022 Jul;54(7):963-975. doi: 10.1038/s41588-022-01100-4. Epub 2022 Jun 30. PMID: 35773407; PMCID: PMC9279158.

16. Ranganathan P, Agrawal A, Bhushan R, Chavalmane AK, Kalathur RK, Takahashi T, Kondaiah P. Expression profiling of genes regulated by TGF-beta: differential regulation in normal and tumour cells. *BMC Genomics*. 2007 Apr 11;8:98. doi: 10.1186/1471-2164-8-98. PMID: 17425807; PMCID: PMC1858692.
17. Kurundkar AR, Kurundkar D, Rangarajan S, Locy ML, Zhou Y, Liu RM, Zmijewski J, Thannickal VJ. The matricellular protein CCN1 enhances TGF- β 1/SMAD3-dependent profibrotic signaling in fibroblasts and contributes to fibrogenic responses to lung injury. *FASEB J*. 2016 Jun;30(6):2135-50. doi: 10.1096/fj.201500173. Epub 2016 Feb 16. PMID: 26884454; PMCID: PMC4871800.
18. Oshi M, Roy AM, Yan L, Kinoshita S, Tamura Y, Kosaka T, Akiyama H, Kunisaki C, Takabe K, Endo I. Enhanced epithelial-mesenchymal transition signatures are linked with adverse tumor microenvironment, angiogenesis and worse survival in gastric cancer. *Cancer Gene Ther*. 2024 May;31(5):746-754. doi: 10.1038/s41417-024-00756-w. Epub 2024 Mar 26. PMID: 38532115.
19. Chu FF, Doroshow JH, Esworthy RS. Expression, characterization, and tissue distribution of a new cellular selenium-dependent glutathione peroxidase, GSHPx-GI. *J Biol Chem*. 1993 Feb 5;268(4):2571-6. PMID: 8428933.
20. Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, Jiang G, Hsiao J, Mermel CH, Getz G, Barretina J, Gopal S, Tamayo P, Gould J, Tsherniak A, Stransky N, Luo B, Ren Y, Drapkin R, Bhatia SN, Mesirov JP, Garraway LA, Meyerson M, Lander ES, Root DE, Hahn WC. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc Natl Acad Sci U S A*. 2011 Jul 26;108(30):12372-7. doi: 10.1073/pnas.1109363108. Epub 2011 Jul 11. PMID: 21746896; PMCID: PMC3145679.
21. Leist M, Raab B, Maurer S, Rösick U, Brigelius-Flohé R. Conventional cell culture media do not adequately supply cells with antioxidants and thus facilitate peroxide-induced genotoxicity. *Free Radic Biol Med*. 1996;21(3):297-306. doi: 10.1016/0891-5849(96)00045-7. PMID: 8855440.
22. Touat-Hamici Z, Bulteau AL, Bianga J, Jean-Jacques H, Szpunar J, Lobinski R, Chavatte L. Selenium-regulated hierarchy of human selenoproteome in cancerous and immortalized cells lines. *Biochim Biophys Acta Gen Subj*. 2018 Nov;1862(11):2493-2505. doi: 10.1016/j.bbagen.2018.04.012. Epub 2018 Apr 13. PMID: 29660373.
23. Parant F, Mure F, Maurin J, Beauvilliers L, Chorfa C, El Jamali C, Ohlmann T, Chavatte L. Selenium Discrepancies in Fetal Bovine Serum: Impact on Cellular Selenoprotein Expression. *Int J Mol Sci*. 2024 Jul 1;25(13):7261. doi: 10.3390/ijms25137261. PMID: 39000368; PMCID: PMC11242189.
24. Esworthy RS, Kim BW, Chow J, Shen B, Doroshow JH, Chu FF. Nox1 causes ileocolitis in mice deficient in glutathione peroxidase-1 and -2. *Free Radic Biol Med*. 2014 Mar;68:315-25. doi: 10.1016/j.freeradbiomed.2013.12.018. Epub 2013 Dec 25. PMID: 24374371; PMCID: PMC3943970.
25. Chu FF, Esworthy RS, Doroshow JH, Grasberger H, Donko A, Leto TL, Gao Q, Shen B. Deficiency in Duox2 activity alleviates ileitis in GPx1- and GPx2-knockout mice without affecting apoptosis incidence in the crypt epithelium. *Redox Biol*. 2017 Apr;11:144-156. doi: 10.1016/j.redox.2016.11.001. Epub 2016 Nov 22. PMID: 27930931; PMCID: PMC5148781.
26. Jones RM, Luo L, Ardita CS, Richardson AN, Kwon YM, Mercante JW, Alam A, Gates CL, Wu H, Swanson PA, Lambeth JD, Denning PW, Neish AS. Symbiotic lactobacilli stimulate gut epithelial proliferation via Nox-mediated generation of reactive oxygen species. *EMBO J*. 2013 Nov 27;32(23):3017-28. doi: 10.1038/emboj.2013.224. Epub 2013 Oct 18. PMID: 24141879; PMCID: PMC3844951.
27. Kato M, Marumo M, Nakayama J, Matsumoto M, Yabe-Nishimura C, Kamata T. The ROS-generating oxidase Nox1 is required for epithelial restitution following colitis. *Exp Anim*. 2016 Jul 29;65(3):197-205. doi: 10.1538/expanim.15-0127. Epub 2016 Feb 15. PMID: 26876598; PMCID: PMC4976233.
28. Matsumoto M, Katsuyama M, Iwata K, Ibi M, Zhang J, Zhu K, Nauseef WM, Yabe-Nishimura C. Characterization of N-glycosylation sites on the extracellular domain of NOX1/NADPH oxidase. *Free Radic Biol Med*. 2014 Mar;68:196-204. doi: 10.1016/j.freeradbiomed.2013.12.013. Epub 2013 Dec 19. PMID: 24361341.
29. Brzozowa-Zasada M, Ianaro A, Piecuch A, Michalski M, Matysiak N, Stęplewska K. Immunohistochemical Expression of Glutathione Peroxidase-2 (Gpx-2) and Its Clinical Relevance in Colon Adenocarcinoma Patients. *Int J Mol Sci*. 2023 Sep 27;24(19):14650. doi: 10.3390/ijms241914650. PMID: 37834097; PMCID: PMC10572251.
30. Komatsu H, Okayasu I, Mitomi H, Imai H, Nakagawa Y, Obata F. Immunohistochemical detection of human gastrointestinal glutathione peroxidase in normal tissues and cultured cells with novel mouse monoclonal antibodies. *J Histochem Cytochem*. 2001 Jun;49(6):759-66. doi: 10.1177/002215540104900609. PMID: 11373322.
31. Florian S, Wingler K, Schmehl K, Jacobasch G, Kreuzer OJ, Meyerhof W, Brigelius-Flohé R. Cellular and subcellular localization of gastrointestinal glutathione peroxidase in normal and malignant human intestinal tissue. *Free Radic Res*. 2001 Dec;35(6):655-63. doi: 10.1080/10715760100301181. PMID: 11811519.

32. Becker WR, Nevins SA, Chen DC, Chiu R, Horning AM, Guha TK, Laquindanum R, Mills M, Chaib H, Ladabaum U, Longacre T, Shen J, Esplin ED, Kundaje A, Ford JM, Curtis C, Snyder MP, Greenleaf WJ. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet.* 2022 Jul;54(7):985-995. doi: 10.1038/s41588-022-01088-x. Epub 2022 Jun 20. PMID: 35726067; PMCID: PMC9279149.
33. Sasaki N, Sachs N, Wiebrands K, Ellenbroek SI, Fumagalli A, Lyubimova A, Begthel H, van den Born M, van Es JH, Karthaus WR, Li VS, López-Iglesias C, Peters PJ, van Rheenen J, van Oudenaarden A, Clevers H. Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. *Proc Natl Acad Sci U S A.* 2016 Sep 13;113(37):E5399-407. doi: 10.1073/pnas.1607327113. Epub 2016 Aug 29. PMID: 27573849; PMCID: PMC5027439.
34. Pilat JM, Brown RE, Chen Z, Berle NJ, Othon AP, Washington MK, Anant SA, Kurokawa S, Ng VH, Thompson JJ, Jacobse J, Goettel JA, Lee E, Choksi YA, Lau KS, Short SP, Williams CS. SELENOP modifies sporadic colorectal carcinogenesis and WNT signaling activity through LRP5/6 interactions. *J Clin Invest.* 2023 Jul 3;133(13):e165988. doi: 10.1172/JCI165988. PMID: 37166989; PMCID: PMC10313376.
35. Schmitt M, Schewe M, Sacchetti A, Feijtel D, van de Geer WS, Teeuwssen M, Sleddens HF, Joosten R, van Royen ME, van de Werken HJG, van Es J, Clevers H, Fodde R. Paneth Cells Respond to Inflammation and Contribute to Tissue Regeneration by Acquiring Stem-like Features through SCF/c-Kit Signaling. *Cell Rep.* 2018 Aug 28;24(9):2312-2328.e7. doi: 10.1016/j.celrep.2018.07.085. PMID: 30157426.
36. Yu S, Tong K, Zhao Y, Balasubramanian I, Yap GS, Ferraris RP, Bonder EM, Verzi MP, Gao N. Paneth Cell Multipotency Induced by Notch Activation following Injury. *Cell Stem Cell.* 2018 Jul 5;23(1):46-59.e5. doi: 10.1016/j.stem.2018.05.002. Epub 2018 Jun 7. PMID: 29887318; PMCID: PMC6035085.
37. Mei X, Gu M, Li M. Plasticity of Paneth cells and their ability to regulate intestinal stem cells. *Stem Cell Res Ther.* 2020 Aug 12;11(1):349. doi: 10.1186/s13287-020-01857-7. PMID: 32787930; PMCID: PMC7425583.
38. Fumagalli A, Oost KC, Kester L, Morgner J, Bornes L, Bruens L, Spaargaren L, Azkanaz M, Schelfhorst T, Beerling E, Heinz MC, Postrach D, Seinstra D, Sieuwerts AM, Martens JWM, van der Elst S, van Baalen M, Bhowmick D, Vrisekoop N, Ellenbroek SIJ, Suijkerbuijk SJE, Snippert HJ, van Rheenen J. Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer. *Cell Stem Cell.* 2020 Apr 2;26(4):569-578.e7. doi: 10.1016/j.stem.2020.02.008. Epub 2020 Mar 12. PMID: 32169167; PMCID: PMC7118369.
39. Hong Y, Liew SC, Thean LF, Tang CL, Cheah PY. Human colorectal cancer initiation is bidirectional, and cell growth, metabolic genes and transporter genes are early drivers of tumorigenesis. *Cancer Lett.* 2018 Sep 1;431:213-218. doi: 10.1016/j.canlet.2018.06.005. Epub 2018 Jun 6. PMID: 29885515.
40. Yuan K, Liu Y, Chen HN, Zhang L, Lan J, Gao W, Dou Q, Nice EC, Huang C. Thiol-based redox proteomics in cancer research. *Proteomics.* 2015 Jan;15(2-3):287-99. doi: 10.1002/pmic.201400164. Epub 2014 Dec 15. PMID: 25251260.
41. Nguyen Huu T, Park J, Zhang Y, Park I, Yoon HJ, Woo HA, Lee SR. Redox Regulation of PTEN by Peroxiredoxins. *Antioxidants (Basel).* 2021 Feb 16;10(2):302. doi: 10.3390/antiox10020302. PMID: 33669370; PMCID: PMC7920247.
42. Esworthy RS, Swiderek KM, Ho YS, Chu FF. Selenium-dependent glutathione peroxidase-GI is a major glutathione peroxidase activity in the mucosal epithelium of rodent intestine. *Biochim Biophys Acta.* 1998 Jul 23;1381(2):213-26. doi: 10.1016/s0304-4165(98)00032-4. PMID: 9685647.
43. Peng F, Xu Q, Jing X, Chi X, Zhang Z, Meng X, Liu X, Yan J, Liu X, Shao S. GPX2 promotes EMT and metastasis in non-small cell lung cancer by activating PI3K/AKT/mTOR/Snail signaling axis. *FASEB Bioadv.* 2023 Apr 8;5(6):233-250. doi: 10.1096/fba.2022-00045. PMID: 37287867; PMCID: PMC10242197.
44. Haider S, Tyekucheva S, Prandi D, Fox NS, Ahn J, Xu AW, Pantazi A, Park PJ, Laird PW, Sander C, Wang W, Demichelis F, Loda M, Boutros PC; Cancer Genome Atlas Research Network. Systematic Assessment of Tumor Purity and Its Clinical Implications. *JCO Precis Oncol.* 2020 Sep 4;4:PO.20.00016. doi: 10.1200/PO.20.00016. PMID: 33015524; PMCID: PMC7529507.
45. Hiller F, Besselt K, Deubel S, Brigelius-Flohé R, Kipp AP. GPx2 Induction Is Mediated Through STAT Transcription Factors During Acute Colitis. *Inflamm Bowel Dis.* 2015 Sep;21(9):2078-89. doi: 10.1097/MIB.0000000000000464. PMID: 26115075
46. Florian S, Krehl S, Loewinger M, Kipp A, Banning A, Esworthy S, Chu FF, Brigelius-Flohé R. Loss of GPx2 increases apoptosis, mitosis, and GPx1 expression in the intestine of mice. *Free Radic Biol Med.* 2010 Dec 1;49(11):1694-702. doi: 10.1016/j.freeradbiomed.2010.08.029. Epub 2010 Sep 7. PMID: 20828612; PMCID: PMC4132893.
47. Mouradov D, Sloggett C, Jorissen RN, Love CG, Li S, Burgess AW, Arango D, Strausberg RL, Buchanan D, Wormald S, O'Connor L, Wilding JL, Bicknell D, Tomlinson IP, Bodmer WF, Mariadason JM, Sieber OM. Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Res.* 2014 Jun 15;74(12):3238-47. doi: 10.1158/0008-5472.CAN-14-0013. Epub 2014 Apr 22. PMID: 24755471.
48. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, Petti C, Fiori A, Orzan F, Senetta R, Boccaccio C, Ficarra E, Marchionni L, Trusolino L, Medico E, Bertotti A. Selective analysis of cancer-cell

- intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat Commun*. 2017 May 31;8:15107. doi: 10.1038/ncomms15107. PMID: 28561063; PMCID: PMC5499209.
49. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Sonesson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015 Nov;21(11):1350-6. doi: 10.1038/nm.3967. Epub 2015 Oct 12. PMID: 26457759; PMCID: PMC4636487.
 50. Ziskin JL, Dunlap D, Yaylaoglu M, Fodor IK, Forrest WF, Patel R, Ge N, Hutchins GG, Pine JK, Quirke P, Koeppen H, Jubb AM. In situ validation of an intestinal stem cell signature in colorectal cancer. *Gut*. 2013 Jul;62(7):1012-23. doi: 10.1136/gutjnl-2011-301195. Epub 2012 May 25. PMID: 22637696
 51. Kipp AP, Müller MF, Göken EM, Deubel S, Brigelius-Flohé R. The selenoproteins GPx2, TrxR2 and TrxR3 are regulated by Wnt signalling in the intestinal epithelium. *Biochim Biophys Acta*. 2012 Oct;1820(10):1588-96. doi: 10.1016/j.bbagen.2012.05.016. Epub 2012 Jun 7. PMID: 22683372.
 52. Ahmed KM, Veeramachaneni R, Deng D, Putluri N, Putluri V, Cardenas MF, Wheeler DA, Decker WK, Frederick AI, Kazi S, Sikora AG, Sandulache VC, Frederick MJ. Glutathione peroxidase 2 is a metabolic driver of the tumor immune microenvironment and immune checkpoint inhibitor response. *J Immunother Cancer*. 2022 Aug;10(8):e004752. doi: 10.1136/jitc-2022-004752. PMID: 36002187; PMCID: PMC9413193.
 53. Smiraglia DJ, Rush LJ, Frühwald MC, Dai Z, Held WA, Costello JF, Lang JC, Eng C, Li B, Wright FA, Caligiuri MA, Plass C. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum Mol Genet*. 2001 Jun 15;10(13):1413-9. doi: 10.1093/hmg/10.13.1413. PMID: 11440994.
 54. Suter CM, Norrie M, Ku SL, Cheong KF, Tomlinson I, Ward RL. CpG island methylation is a common finding in colorectal cancer cell lines. *Br J Cancer*. 2003 Feb 10;88(3):413-9. doi: 10.1038/sj.bjc.6600699. PMID: 12569385; PMCID: PMC2747532.
 55. Auman JT, McLeod HL. Colorectal cancer cell lines lack the molecular heterogeneity of clinical colorectal tumors. *Clin Colorectal Cancer*. 2010 Jan;9(1):40-7. doi: 10.3816/CCC.2010.n.005. PMID: 20100687.
 56. Emmink BL, Laoukili J, Kipp AP, Koster J, Govaert KM, Fatrai S, Verheem A, Steller EJ, Brigelius-Flohé R, Jimenez CR, Borel Rinkes IH, Kranenburg O. GPx2 suppression of H2O2 stress links the formation of differentiated tumor mass to metastatic capacity in colorectal cancer. *Cancer Res*. 2014 Nov 15;74(22):6717-30. doi: 10.1158/0008-5472.CAN-14-1645. Epub 2014 Sep 26. PMID: 25261240.
 57. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, Campbell K, Cano A, Casanova J, Christofori G, Dedhar S, Derynck R, Ford HL, Fuxe J, García de Herreros A, Goodall GJ, Hadjantonakis AK, Huang RYJ, Kalchauer C, Kalluri R, Kang Y, Khew-Goodall Y, Levine H, Liu J, Longmore GD, Mani SA, Massagué J, Mayor R, McClay D, Mostov KE, Newgreen DF, Nieto MA, Puisieux A, Runyan R, Savagner P, Stanger B, Stemmler MP, Takahashi Y, Takeichi M, Theveneau E, Thiery JP, Thompson EW, Weinberg RA, Williams ED, Xing J, Zhou BP, Sheng G; EMT International Association (TEMTIA). Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. 2020 Jun;21(6):341-352. doi: 10.1038/s41580-020-0237-9. Epub 2020 Apr 16. Erratum in: *Nat Rev Mol Cell Biol*. 2021 Dec;22(12):834. PMID: 32300252; PMCID: PMC7250738.
 58. Loeser H, Scholz M, Fuchs H, Essakly A, Damanakis AI, Zander T, Büttner R, Schröder W, Bruns C, Quaas A, Gebauer F. Integrin alpha V (ITGAV) expression in esophageal adenocarcinoma is associated with shortened overall-survival. *Sci Rep*. 2020 Oct 27;10(1):18411. doi: 10.1038/s41598-020-75085-7. PMID: 33110104; PMCID: PMC7591891.
 59. Kurundkar AR, Kurundkar D, Rangarajan S, Locy ML, Zhou Y, Liu RM, Zmijewski J, Thannickal VJ. The matricellular protein CCN1 enhances TGF- β 1/SMAD3-dependent profibrotic signaling in fibroblasts and contributes to fibrogenic responses to lung injury. *FASEB J*. 2016 Jun;30(6):2135-50. doi: 10.1096/fj.201500173. Epub 2016 Feb 16. PMID: 26884454; PMCID: PMC4871800.
 60. Huet C, Sahuquillo-Merino C, Coudrier E, Louvard D. Absorptive and mucus-secreting subclones isolated from a multipotent intestinal cell line (HT-29) provide new models for cell polarity and terminal differentiation. *J Cell Biol*. 1987 Jul;105(1):345-57. doi: 10.1083/jcb.105.1.345. PMID: 3611191; PMCID: PMC211493
 61. Bravard A, Beaumatin J, Dussaulx E, Lesuffleur T, Zweibaum A, Luccioni C. Modifications of the antioxidant metabolism during proliferation and differentiation of colon tumor cell lines. *Int J Cancer*. 1994 Dec 15;59(6):843-7. doi: 10.1002/ijc.2910590622. PMID: 7989127
 62. Godtliebsen G, Larsen KB, Bhujabal Z, Opstad IS, Nager M, Punnakal AR, Kalstad TB, Olsen R, Lund T, Prasad DK, Agarwal K, Myrnes T, Birgisdottir AB. High-resolution visualization and assessment of basal and OXPHOS-induced mitophagy in H9c2 cardiomyoblasts. *Autophagy*. 2023 Oct;19(10):2769-2788. doi: 10.1080/15548627.2023.2230837. Epub 2023 Jul 5. PMID: 37405374; PMCID: PMC10472865
 63. Hewitt RE, McMarlin A, Kleiner D, Wersto R, Martin P, Tsokos M, Stamp GW, Stetler-Stevenson WG. Validation of a model of colon cancer progression. *J Pathol*. 2000 Dec;192(4):446-54. doi: 10.1002/1096-

- 9896(2000)9999:9999<::AID-PATH775>3.0.CO;2-K. Erratum in: J Pathol 2001 Aug;194(4):507. Tsoskas M [corrected to Tsokos M]. PMID: 11113861.
64. Xu Y, Zhang L, Wang Q, Zheng M. Comparison of Different Colorectal Cancer With Liver Metastases Models Using Six Colorectal Cancer Cell Lines. *Pathol Oncol Res.* 2020 Oct;26(4):2177-2183. doi: 10.1007/s12253-020-00805-3. Epub 2020 Mar 14. PMID: 32172478.
 65. Sun MA, Wang Y, Cheng H, Zhang Q, Ge W, Guo D. RedoxDB--a curated database for experimentally verified protein oxidative modification. *Bioinformatics.* 2012 Oct 1;28(19):2551-2. doi: 10.1093/bioinformatics/bts468. Epub 2012 Jul 25. PMID: 22833525.
 66. Merkley ED, Metz TO, Smith RD, Baynes JW, Frizzell N. The succinated proteome. *Mass Spectrom Rev.* 2014 Mar-Apr;33(2):98-109. doi: 10.1002/mas.21382. Epub 2013 Sep 30. PMID: 24115015; PMCID: PMC4038156.
 67. Dustin CM, Heppner DE, Lin MJ, van der Vliet A. Redox regulation of tyrosine kinase signalling: more than meets the eye. *J Biochem.* 2020 Feb 1;167(2):151-163. doi: 10.1093/jb/mvz085. PMID: 31599960; PMCID: PMC6988750.
 68. Welsh CL, Madan LK. Protein Tyrosine Phosphatase regulation by Reactive Oxygen Species. *Adv Cancer Res.* 2024;162:45-74. doi: 10.1016/bs.acr.2024.05.002. Epub 2024 May 24. PMID: 39069369.]
 69. Han L, Wang S, Wei C, Fang Y, Huang S, Yin T, Xiong B, Yang C. Tumour microenvironment: a non-negligible driver for epithelial-mesenchymal transition in colorectal cancer. *Expert Rev Mol Med.* 2021 Nov 11;23:e16. doi: 10.1017/erm.2021.13. PMID: 34758892.
 70. Catalano V, Turdo A, Di Franco S, Dieli F, Todaro M, Stassi G. Tumor and its microenvironment: a synergistic interplay. *Semin Cancer Biol.* 2013 Dec;23(6 Pt B):522-32. doi: 10.1016/j.semcancer.2013.08.007. Epub 2013 Sep 4. PMID: 24012661.
 71. Ernst M, Ramsay RG. Colorectal cancer mouse models: integrating inflammation and the stroma. *J Gastroenterol Hepatol.* 2012 Jan;27(1):39-50. doi: 10.1111/j.1440-1746.2011.06883.x. PMID: 22188027.
 72. Hirsch D, Seyfried S, Staib T, Fiedler D, Sauer C, Ried T, Witt S, Rueckert F, Gaiser T. Newly established gastrointestinal cancer cell lines retain the genomic and immunophenotypic landscape of their parental cancers. *Sci Rep.* 2020 Oct 21;10(1):17895. doi: 10.1038/s41598-020-74797-0. PMID: 33087752; PMCID: PMC7578805.
 73. Tompkins WA, Watrach AM, Schmale JD, Schultz RM, Harris JA. Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum. *J Natl Cancer Inst.* 1974 Apr;52(4):1101-10. doi: 10.1093/jnci/52.4.1101. PMID: 4826581.
 74. Calon A, Espinet E, Palomo-Ponce S, Tauriello DV, Iglesias M, Céspedes MV, Sevillano M, Nadal C, Jung P, Zhang XH, Byrom D, Riera A, Rossell D, Mangués R, Massagué J, Sancho E, Batlle E. Dependency of colorectal cancer on a TGF- β -driven program in stromal cells for metastasis initiation. *Cancer Cell.* 2012 Nov 13;22(5):571-84. doi: 10.1016/j.ccr.2012.08.013. PMID: 23153532; PMCID: PMC3512565
 75. Fagundes RR, Belt SC, Bakker BM, Dijkstra G, Harmsen HJM, Faber KN. Beyond butyrate: microbial fiber metabolism supporting colonic epithelial homeostasis. *Trends Microbiol.* 2024 Feb;32(2):178-189. doi: 10.1016/j.tim.2023.07.014. Epub 2023 Aug 16. PMID: 37596118
 76. Tomé D. The Roles of Dietary Glutamate in the Intestine. *Ann Nutr Metab.* 2018;73 Suppl 5:15-20. doi: 10.1159/000494777. Epub 2018 Dec 3. PMID: 30508814
 77. Gomes S, Baltazar F, Silva E, Preto A. Microbiota-Derived Short-Chain Fatty Acids: New Road in Colorectal Cancer Therapy. *Pharmaceutics.* 2022 Nov 1;14(11):2359. doi: 10.3390/pharmaceutics14112359. PMID: 36365177; PMCID: PMC9698921
 78. Teeli AS, Łuczyńska K, Haque E, Gayas MA, Winiarczyk D, Taniguchi H. Disruption of Tumor Suppressors HNF4 α /HNF1 α Causes Tumorigenesis in Liver. *Cancers (Basel).* 2021 Oct 26;13(21):5357. doi: 10.3390/cancers13215357. PMID: 34771521; PMCID: PMC8582545.
 79. D'Angelo A, Bluteau O, Garcia-Gonzalez MA, Gresh L, Doyen A, Garbay S, Robine S, Pontoglio M. Hepatocyte nuclear factor 1 α and beta control terminal differentiation and cell fate commitment in the gut epithelium. *Development.* 2010 May;137(9):1573-82. doi: 10.1242/dev.044420. PMID: 20388655
 80. Qu X, Yan X, Kong C, Zhu Y, Li H, Pan D, Zhang X, Liu Y, Yin F, Qin H. c-Myb promotes growth and metastasis of colorectal cancer through c-fos-induced epithelial-mesenchymal transition. *Cancer Sci.* 2019 Oct;110(10):3183-3196. doi: 10.1111/cas.14141. Epub 2019 Aug 13. PMID: 31338937; PMCID: PMC6778643.
 81. Fu M, Hu Y, Lan T, Guan KL, Luo T, Luo M. The Hippo signalling pathway and its implications in human health and diseases. *Signal Transduct Target Ther.* 2022 Nov 8;7(1):376. doi: 10.1038/s41392-022-01191-9
 82. Huang H, Zhang W, Pan Y, Gao Y, Deng L, Li F, Li F, Ma X, Hou S, Xu J, Li P, Li X, Hu G, Li C, Chen H, Zhang L, Ji H. YAP Suppresses Lung Squamous Cell Carcinoma Progression via Deregulation of the DNp63-GPX2 Axis and ROS Accumulation. *Cancer Res.* 2017 Nov 1;77(21):5769-5781. doi: 10.1158/0008-5472.CAN-17-0449. Epub 2017 Sep 15.
 83. Soo HC, Chung FF, Lim KH, Yap VA, Bradshaw TD, Hii LW, Tan SH, See SJ, Tan YF, Leong CO, Mai CW. Cudraflavone C Induces Tumor-Specific Apoptosis in Colorectal Cancer Cells through Inhibition of the

- Phosphoinositide 3-Kinase (PI3K)-AKT Pathway. *PLoS One*. 2017 Jan 20;12(1):e0170551. doi: 10.1371/journal.pone.0170551. PMID: 28107519; PMCID: PMC5249192.
84. Coant N, Ben Mkaddem S, Pedruzzi E, Guichard C, Tréton X, Ducroc R, Freund JN, Cazals-Hatem D, Bouhnik Y, Woerther PL, Skurnik D, Grodet A, Fay M, Biard D, Lesuffleur T, Deffert C, Moreau R, Groyer A, Krause KH, Daniel F, Ogier-Denis E. NADPH oxidase 1 modulates WNT and NOTCH1 signaling to control the fate of proliferative progenitor cells in the colon. *Mol Cell Biol*. 2010 Jun;30(11):2636-50. doi: 10.1128/MCB.01194-09. Epub 2010 Mar 29. PMID: 20351171; PMCID: PMC2876517.
 85. Chu FF, Esworthy RS, Shen B, Doroshov JH. Role of the microbiota in ileitis of a mouse model of inflammatory bowel disease-Glutathione peroxidase 1 and 2-double knockout mice on a C57BL background. *Microbiologyopen*. 2020 Oct;9(10):e1107. doi: 10.1002/mbo3.1107. Epub 2020 Aug 18. PMID: 32810389; PMCID: PMC7568258.
 86. Chu FF, Esworthy RS, Chu PG, Longmate JA, Huycke MM, Wilczynski S, Doroshov JH. Bacteria-induced intestinal cancer in mice with disrupted Gpx1 and Gpx2 genes. *Cancer Res*. 2004 Feb 1;64(3):962-8. doi: 10.1158/0008-5472.can-03-2272. PMID: 14871826.
 87. Lu J, Jiang G, Wu Y, Antony S, Meitzler JL, Juhasz A, Liu H, Roy K, Makhoul H, Chuaqui R, Butcher D, Konaté MM, Doroshov JH. NADPH oxidase 1 is highly expressed in human large and small bowel cancers. *PLoS One*. 2020 May 19;15(5):e0233208. doi: 10.1371/journal.pone.0233208. PMID: 32428030; PMCID: PMC7237001.
 88. Dunne PD, McArt DG, Bradley CA, O'Reilly PG, Barrett HL, Cummins R, O'Grady T, Arthur K, Loughrey MB, Allen WL, McDade SS, Waugh DJ, Hamilton PW, Longley DB, Kay EW, Johnston PG, Lawler M, Salto-Tellez M, Van Schaeysbroeck S. Challenging the Cancer Molecular Stratification Dogma: Intratumoral Heterogeneity Undermines Consensus Molecular Subtypes and Potential Diagnostic Value in Colorectal Cancer. *Clin Cancer Res*. 2016 Aug 15;22(16):4095-104. doi: 10.1158/1078-0432.CCR-16-0032. Epub 2016 May 5. PMID: 27151745
 89. Leibovitz A, Stinson JC, McCombs WB 3rd, McCoy CE, Mazur KC, Mabry ND. Classification of human colorectal adenocarcinoma cell lines. *Cancer Res*. 1976 Dec;36(12):4562-9. PMID: 1000501.
 90. Solic N, Collins JE, Richter A, Holt SJ, Campbell I, Alexander P, Davies DE. Two newly established cell lines derived from the same colonic adenocarcinoma exhibit differences in EGF-receptor ligand and adhesion molecule expression. *Int J Cancer*. 1995 Jul 4;62(1):48-57. doi: 10.1002/ijc.2910620111. PMID: 7601566.
 91. Ohata H, Shiokawa D, Obata Y, Sato A, Sakai H, Fukami M, Hara W, Taniguchi H, Ono M, Nakagama H, Okamoto K. NOX1-Dependent mTORC1 Activation via S100A9 Oxidation in Cancer Stem-like Cells Leads to Colon Cancer Progression. *Cell Rep*. 2019 Jul 30;28(5):1282-1295.e8. doi: 10.1016/j.celrep.2019.06.085. PMID: 31365870.
 92. van der Post S, Birchenough GMH, Held JM. NOX1-dependent redox signaling potentiates colonic stem cell proliferation to adapt to the intestinal microbiota by linking EGFR and TLR activation. *Cell Rep*. 2021 Apr 6;35(1):108949. doi: 10.1016/j.celrep.2021.108949. PMID: 33826887; PMCID: PMC10327654.
 93. Chen M, Mainardi S, Liefstink C, Velds A, de Rink I, Yang C, Kuiken HJ, Morris B, Edwards F, Jochems F, van Tellingen O, Boeijs M, Proost N, Jansen RA, Qin S, Jin H, Koen van der Mijl JC, Schepers A, Venkatesan S, Qin W, Beijersbergen RL, Wang L, Bernards R. Targeting of vulnerabilities of drug-tolerant persisters identified through functional genetics delays tumor relapse. *Cell Rep Med*. 2024 Mar 19;5(3):101471. doi: 10.1016/j.xcrm.2024.101471. PMID: 38508142; PMCID: PMC10983104.
 94. Oren Y, Tsabar M, Cuoco MS, Amir-Zilberstein L, Cabanos HF, Hütter JC, Hu B, Thakore PI, Tabaka M, Fulco CP, Colgan W, Cuevas BM, Hurvitz SA, Slamon DJ, Deik A, Pierce KA, Clish C, Hata AN, Zaganjor E, Lahav G, Politi K, Brugge JS, Regev A. Cycling cancer persister cells arise from lineages with distinct programs. *Nature*. 2021 Aug;596(7873):576-582. doi: 10.1038/s41586-021-03796-6. Epub 2021 Aug 11. PMID: 34381210; PMCID: PMC9209846.
 95. Barry CJ, Pillay CS, Rohwer JM. Modelling the Decamerisation Cycle of PRDX1 and the Inhibition-like Effect on Its Peroxidase Activity. *Antioxidants (Basel)*. 2023 Sep 1;12(9):1707. doi: 10.3390/antiox12091707. PMID: 37760010; PMCID: PMC10525498.
 96. Portillo-Ledesma S, Randall LM, Parsonage D, Dalla Rizza J, Karplus PA, Poole LB, Denicola A, Ferrer-Sueta G. Differential Kinetics of Two-Cysteine Peroxiredoxin Disulfide Formation Reveal a Novel Model for Peroxide Sensing. *Biochemistry*. 2018 Jun 19;57(24):3416-3424. doi: 10.1021/acs.biochem.8b00188. Epub 2018 Mar 30. PMID: 29553725; PMCID: PMC6383210.
 97. Low FM, Hampton MB, Peskin AV, Winterbourn CC. Peroxiredoxin 2 functions as a noncatalytic scavenger of low-level hydrogen peroxide in the erythrocyte. *Blood*. 2007 Mar 15;109(6):2611-7. doi: 10.1182/blood-2006-09-048728. Epub 2006 Nov 14. PMID: 17105810.
 98. Buetler TM, Krauskopf A, Ruegg UT. Role of superoxide as a signaling molecule. *News Physiol Sci*. 2004 Jun;19:120-3. doi: 10.1152/nips.01514.2003. PMID: 15143206.

99. Barbáchano A, Fernández-Barral A, Bustamante-Madrid P, Prieto I, Rodríguez-Salas N, Larriba MJ, Muñoz A. Organoids and Colorectal Cancer. *Cancers (Basel)*. 2021 May 28;13(11):2657. doi: 10.3390/cancers13112657. PMID: 34071313; PMCID: PMC8197877
100. Müller MF, Florian S, Pommer S, Osterhoff M, Esworthy RS, Chu FF, Brigelius-Flohé R, Kipp AP. Deletion of glutathione peroxidase-2 inhibits azoxymethane-induced colon cancer development. *PLoS One*. 2013 Aug 19;8(8):e72055. doi: 10.1371/journal.pone.0072055. PMID: 23977205; PMCID: PMC3747154.
101. Banning A, Kipp A, Schmitmeier S, Löwinger M, Florian S, Krehl S, Thalmann S, Thierbach R, Steinberg P, Brigelius-Flohé R. Glutathione Peroxidase 2 Inhibits Cyclooxygenase-2-Mediated Migration and Invasion of HT-29 Adenocarcinoma Cells but Supports Their Growth as Tumors in Nude Mice. *Cancer Res*. 2008 Dec 1;68(23):9746-53. doi: 10.1158/0008-5472.CAN-08-1321. PMID: 19047153.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.