Reyes (2020)

Supplementary Material and Methods

Clinical Trial. A total of 80 patients were recruited to the clinical trial (MDACC IRB# 2013-0698) at four academic cancer centers (The University of Texas MD Anderson Cancer Center, Dana-Farber/Brigham and Women's Cancer Center, University of Michigan, Huntsman Cancer Institute). This study was sponsored by the National Cancer Institute (NCI) and was initially approved by the Institutional Review Board (IRB) of The University of Texas MD Anderson Cancer Center (MDACC) on 01/27/2014. Participants were identified and recruited from the clinical cancer genetics and gastroenterology research registries.

Eligible participants were men and women, ≥18 years of age diagnosed with Lynch Syndrome (LS) or Lynch-like Syndrome (LLS) defined as meeting any of the following: 1. Carriers or obligate carriers (by pedigree) of a pathogenic mutation in one of the four DNA mismatch repair (MMR) genes (i.e. *MLH1, MSH2/EPCAM, MSH6,* or *PMS2*); 2. patients with a personal history of a non-sporadic MMR deficient premalignant colorectal lesion or tumor, where "non-sporadic MMR deficient" was defined by the presence of MMR deficiency by either loss of expression of any of the MMR proteins by immunohistochemistry o high levels of microsatellite instability (MSI) assessed by MSI testing or both, but without evidence of *MLH1* promoter methylation and/or *BRAF*. This group of patients could have undergone germline MMR genetic testing showing either a variant of unknown significance, absence of pathogenic mutation or had declined germline MMR genetic testing. In addition, all participants had to be free of active/recurrent malignant disease in the 6 months previous to the enrollment, and at least 6 months prior did not have any form of cancer-directed treatment; therefore, survivors and healthy carriers were both potentially eligible for the study, thus exploring the effects of naproxen for primary and secondary cancer prevention. Finally, participants were required to discontinue any NSAID or COX-Inhibitor during the trial.

Reyes (2020)

Exclusion criteria included the following: unwillingness to discontinue taking NSAIDs for the duration of the trial (including cardioprotective aspirin), absence of the use of effective birth control in women of childbearing age, pregnancy or breastfeeding, a white blood cell count of less than 3000/μL, a platelet count of less than 100×10^3/μL, a hemoglobin level of less than 10 g/dL, a serum creatinine level of more than 1.5 mg/dL, transaminases/bilirubin elevations 1.5- to 2-fold above the upper limit of normal, symptoms or features of active gastrointestinal bleeding, cardiovascular events in the previous 5 years, history of allergy or hypersensitivity to naproxen, history of cancer within the past 6 months, prior proctocolectomy, or prior cancer treatment within the preceding 6 months.

Following informed consent, baseline testing, and pre-study evaluations participants were registered into the database of record with a unique Patient ID (PID) number. Baseline medical history, symptom assessment, use of concomitant medications, a physical examination and a serum pregnancy test in women of childbearing potential were completed prior to the endoscopy procedure. At the completion of the screening, eligibility was confirmed, and participants were randomized to receive placebo (microcrystalline cellulose NF, Avicel PH-102), naproxen sodium at 220 mg p.o. daily, or 440 mg p.o. daily for a total of 6 months. The randomization was done in blocks of 3 or 6. Upon enrollment, each participant was assigned a randomization number that corresponded to a treatment on a randomization list available only to the unblinded study pharmacist. Naproxen (FDA IND exempted) and identically appearing placebo tablets were provided by the NCI Division of Cancer Prevention.

Colorectal mucosa biopsies and associated assessments. Participants were scheduled for a standard of care lower GI endoscopy (flexible sigmoidoscopy or colonoscopy) with acquisition of a total of 15 jumbo tissue biopsies from the macroscopic normal colorectal mucosa of the rectosigmoid area. This requirement was enforced to limit the variability secondary to regional differences between the right and left side of the colon. In addition, collection of urine and blood was performed. Study participants were contacted within two weeks of initiating drug treatment, and then after 3 and 5.5 months to assess for

Reyes (2020)

potential side effects. After 6 months, study participants underwent a research flexible sigmoidoscopy with the same number of biopsies, urine, and blood collection. Participants returned the remainder of the study agent, so a pill count and a review of the agent diary could be performed.

In regards to the tissue biopsies, biopsy #1 was processed in all participants following standard pathology procedures to confirm the absence of neoplastic or dysplastic lesions in the area being biopsied (M.W.T.), and, in a subgroup of 48 participants, to perform a histomorphometry analysis that included the quantification of the number of intraepithelial lymphocytes (IELs) (surface and crypts of Lieberkühn), and also the assessment of the number and area of mucosa associated lymphoid tissue (MALT). In brief, a total of 48 formalin-fixed paraffin-embedded (FFPE) colorectal biopsy specimens were evaluated (30 pre-treatment and 18 post-treatment tissue samples). Note that these were specimens acquired from participants in only one of the four study sites of the trial. From the total of 48, two biopsy blocks were excluded due to artifacts and insufficient tissue for analysis. Two pathologists (L.S. and A.F.C.) evaluated the H&E slides, determined the quality of these sample, and performed the histomorphometric analysis. Slides were scanned using Aperio AT2 slide scanner (Leica Biosystems) at 20x objective magnification and then the resulting images were analyzed using HALO software v2.3.20089.52 (IndicaLab). This allowed automatic segmentation of the mucosa and epithelium, and the separation of the lamina propria compartment (Tissue Classifier module). The MALT compartment was manually segmented. Glass area, muscularis mucosae, and tissue artifacts were excluded for analysis. Then, automated cell number quantification was obtained per compartment (CytoNuclear module, v1.6) in order to calculate the epithelium and MALT cell densities (n/mm2, N=18 specimens). The IELs had to be manually quantified (Manual Click Counter Tool) to calculate the IELs cell density in the epithelial compartment and the IELs/epithelial cell ratio (N=46 specimens).

Biopsies #2, #3 and #4 were flash frozen and used for the determination of levels of prostaglandin E2 (PGE2) and other prostaglandins (PGs); biopsies #5 and #6 were flash frozen and sent for determination

of naproxen levels; biopsies #7 and #8 were kept in RNALater for mRNA extraction and sequencing studies; the rest of the biopsies were either flash frozen or kept in RNALater and stored centrally at MDACC for future analyses.

Statistical analyses of clinical trial data. Participants' demographic characteristics have been summarized by descriptive statistics including mean, standard deviation (std), median, and range for continuous and frequency and proportion for categorical variables. The distributions of these characteristics for the entire cohort and the evaluable cohort have been compared among treatment arms by Wilcoxon rank sum test and chi-square/Fisher's exact test for continuous and categorical variables, respectively. The primary objective of this trial was evaluated by applying Wilcoxon rank sum test to compare the change in concentration of PGE2 levels in normal colorectal mucosa from baseline to 6 months after treatment administration among the treatment arms (placebo vs low dose naproxen, placebo vs high dose naproxen, and low dose vs high dose naproxen). Initially, power calculations were performed for pairwise comparisons using two-sample t-test. However, upon completion of the trial it was evident that the PGE2, PGM-E and PGs data distribution was skewed and therefore we reasoned that a non-parametric test was more appropriate in this context. Bonferroni correction was considered for multiple comparison adjustment (to maintain the overall nominal significance level of 0.05, p-values are compared to 0.05/3=0.0167). PGE2 concentration level was summarized by mean and std at baseline and post-treatment. Response to treatment is defined as a 30% reduction in PGE2 level from pre-treatment to post-treatment. The point estimate and the 95% exact confidence interval for the response rate was calculated in each arm. Each patient's biomarker level at baseline, post-treatment and its change was plotted by treatment arm to evaluate the biomarker's distribution at each arm. Mean values of biomarkers and their standard error estimates are plotted as a function of time. Naproxen concentration in plasma and normal colorectal mucosa have been summarized by mean, std, median, and range for patients taking naproxen 220 mg, 440mg, or placebo after 6 months of therapy and compared using a two-sample t-test. The change in PG concentration level is compared between treatment arms by applying Wilcoxon rank

sum test. The percentage of change in PGE was defined as 100*(PGE (Post) - PGE (BL))/ PGE (BL) and Pearson correlation was estimated between the percentage of change with naproxen level.

Quantification of PGs in tissue samples. Concentrations of PGs in tissue and urine were performed in the Eicosanoid Core Laboratory at Vanderbilt University Medical Center (supervised by GLM). Each tissue biopsy (#2, 3 and 4) was weighed and subsequently ground up manually with a mortar and pestle in the presence of 1mL of an ice cold methanolic solution containing the cyclooxygenase (COX) inhibitor, indomethacin. The grounded tissue was added to 5mL of ice cold methanol containing 1.0 ng of each of the following internal standards: [2H4]-15-F2t-IsoP ([2H4]-8-iso-PGF2α), [2H4]-PGD2, [2H4]-PGE2, [2H4]-11-dehydro-thromboxane B2 (11-dehydro-TxB2), and [2H4]-6-keto-PGF1α (Cayman Chemicals, Ann Arbor, MI). Samples were vortexed to extract the lipids and then centrifuged for 10 minutes at 5000g to remove solid particulates. The liquid layer was transferred to another tube and the methanol dried under a stream of nitrogen. The residue was reconstituted in 10 mL of 0.01N HCl and then applied to a C-18 Sep-Pak column (Waters, Milford, MA) prewashed with 5 mL methanol and 5 mL H2O (pH 3). The cartridge was washed with 10 mL 0.01N HCl, followed by 10 mL heptane, and compounds were then eluted with 10 mL ethyl acetate:heptane (50:50, v/v). The eluate was evaporated under a stream of dry nitrogen. The residue then was taken up in 75 uL acetonitrile for derivatization with methoxyamine HCl. 300 uL of a 3% solution of methoxyamine HCl in water was added to the solution and allowed to react for 30 minutes. Analytes were next extracted from solution with 1 mL of ethyl acetate. The ethyl acetate was evaporated and compounds are converted to the pentafluorobenzyl (PFB) esters by the addition of 40 µL of a 10% solution of pentafluorobenzyl bromide in acetonitrile and 20 uL of a solution of 10% di-isopropylethanolamine in acetonitrile and allowed to incubate for 30 min at 37ºC. Reagents were dried under nitrogen and the residue reconstituted in 30 µL chloroform and 20 µL methanol and chromatographed on a silica TLC plate to the top of the plate in a solvent system of ethyl acetate:methanole (98:2, v/v). The methyl esters of PGF2α, PGE2, and PGD2 were chromatographed on a separate plate as standards and visualized with 10% phosphomolybdic acid in ethanol by heating. The Rf

Reyes (2020)

values of these methyl ester standards in this solvent system are 0.12, 0.25, and 0.50, respectively. The analyte PGD2 was removed from the TLC plates by scraping the region from 4mm to 2 cm above the PGD2-methyl ester. The analyte was then extracted from the silica with 1 mL ethyl acetate and dried under nitrogen. The analytes PGE2 and TxB2 were removed from the TLC plates by scraping the region from 0.5cm above the PGE2-methyl ester standard to 4mm above the PGD2-methyl ester standard. The analytes were then extracted from the silica with 1 mL ethyl acetate and dried under nitrogen. Finally, PGF2α and 6-keto-PGF1α were removed from the TLC plates by scraping the region from the mid-point of the PGF2α-methyl ester standard to the mid-point of the PGE2-methyl ester standard. The analytes were then extracted from the silica with 1 mL ethyl acetate and dried under nitrogen. At this point there were three separate samples for analysis of PGD2, PGE2 and TxB2, and PGF2 and 6-keto-PGF1 in each of the original starting samples. Following TLC purification, compounds were converted to trimethylsilyl (TMS) ether derivatives by addition of 20 uL N,O-bis(trimethylsilyl)trifluoroacetamide and 10 uL dimethylformamide. The samples were incubated at 37°C for 10 minutes and then dried under nitrogen. The residue was re-dissolved for GC/MS analysis in 10 uL undecane that has been stored over a bed of calcium hydride. GC/NICI-MS was carried out on an Agilent 5973 Inert Mass Selective Detector that was coupled with an Agilent 6890n Network GC system that was interfaced with an Agilent computer. The GC was performed using a 15 m, 0.25 mm film thickness, DB-1701 fused silica capillary column (J and W Scientific, Folsom CA). The column temperature is programmed from 190° to 300°C at 20°C per minute. The major ions generated in the NICI mass spectrum of the PFB ester, TMS ether derivatives of the individual analytes are shown in this table. Levels of endogenous eicosanoids in the biological samples were calculated from the ratio of intensities of the [2H0]- and [2H4]-ions. The inter-day variability for each assay is < 10%. The precision for each assay is ±5% while the accuracy for each assay is 95%.

Analyte	[2H0]-Ion	[2H3] or [2H4]-Ion
PGE2 & PGD2	524	528
-keto-PGF1α	614	618
TxB2	614	618
PGF2α	569	573

Reyes (2020)

Quantification of PGE-M in urine samples. The methodology has been previously described[1 2]. In brief, urine was treated using methoximine HCl (16% w/v) in a 1.5 M sodium acetate solution, diluted in water, and adjusted to pH 3 with acetic acid. Samples were purified by extraction using a C18 SepPak, after which 12.4 ng of O- [2H6]methyloxime PGE-M internal standard was added. Urinary creatinine levels were measured using a test kit from Enzo Life Sciences. The urinary metabolite levels in each sample are normalized using the urinary creatinine level of the sample and expressed in ng/mg creatinine.

Quantification of Naproxen Levels in Normal Colorectal Tissue and Plasma. Levels of naproxen were analyzed in plasma and colorectal mucosa biopsies #5 and #6 by the Eicosanoid Core Laboratory at Vanderbilt University Medical Center (supervised by L.J.M.) by modifying a method previously described[3]. In brief, samples were removed from -80C storage, colonic tissue was homogenized in a solution of methanol and phosphate buffered saline (PBS) and, plasma was aliquoted to 12x75 mm borosilicate test tubes. Naproxen-d3 was used as an internal standard. Samples were purified by SPE on a Phenomenex Strata C18 SPE cartridge as described by Vinci. The reconstituted samples were injected onto a RP LC-MS system, configured in a manner similar described by Vinci. Quantification of naproxen in the submitted plasma samples and tissue specimens was accomplished by comparing them against a standard curve that was prepared from blank human samples spiked with varying amounts of standard naproxen and processed in parallel with blinded samples.

mRNA-Seq Profiling Analysis of Normal Colorectal Tissue and Bioinformatics Analyses. Biopsy specimens #7 and #8 were preserved in RNA-later, thawed on ice, and weighed for analysis. Samples were homogenized in TriZol using 5mm steel beads. We used the PureLink RNA Mini Kit (Invitrogen, San Diego, CA) according to manufacturer's instructions. A DNA digestion step was also performed using PureLink DNAse kit. RNA was eluted in nuclease free water and, and concentration was measured using NanoDrop. Samples were stored in -80C. QC, Poly-A selection, and mRNA library prep was

Reyes (2020)

performed at the Genomics Core at Weill Cornell. A total of 106 samples were sequenced using HiSeq4000 distributed in 15 samples per lane and generating 75 bp paired-end reads.

In regards to the bioinformatics analysis, first sequencing quality was assessed in order to remove low-quality reads using *fastp*[4]. The remaining high-quality reads were aligned to ENSEMBLE GRCh38 human genome reference using *STAR_2.4.0e*[5]. Gene level counts according to ENSEMBLE GRCh38 gene annotation were quantified using *htseq-count*[6]. We decided to exclude genes located on Y chromosome from downstream analyses after observing strong clustering patterns driven by gender.

Differential expression analyses of tissue specimens coming from patients allocated to placebo (Pl), lose-dose (LD) and high-dose (HD) naproxen groups were performed separately using Bioconductor R package *DESeq2*[7], in which a paired sample design was specified for comparing post- to pre-treatment samples. To explore and visualize the relative distribution between samples in relationship to clinical/biological variables of interest in a two-dimensional space, Principal Component Analysis (PCA) was performed using data normalized by Variance Stabilizing Transformation (VST). In LD and HD groups, significant genes with Benjamini-Hochberg (BH)-adjusted *P-value*≤0.05 and absolute value of log2-Fold Change≥0.5 were highlighted and annotated with pathways of interest in volcano plots. Their expression patterns were displayed in an unsupervised hierarchical clustering heat map using Bioconductor R package *ComplexHeatmap*[8].

For Gene Set Enrichment Analysis (GSEA), genes were pre-ranked by log2-Fold Change and analyzed with Bioconductor R package *fgsea*[9], separately on HALLMARK, KEGG, and REACTOME gene sets obtained from the Broad Institute Molecular Signatures Database v6.0, as well as Colorectal Cancer (CRC) related pathway gene sets obtained from Guinney et al[10]. Significant LD- and HD-associated pathways were selected based on BH-adjusted *P-value*≤0.05 in their respective naproxen group and also

BH-adjusted *P-value*>0.05 in PL. Based on Normalized Enrichment Score (NES), the selected pathways with common (NES are both positive or both negative) and distinct (NES are one positive, the other negative) effects in LD and HD naproxen were comparatively displayed. We also performed pathway enrichment analysis combining together the LD and HD data sets (combined naproxen analysis).

For immune cell gene enrichment analysis, gene signatures were obtained from Nirmal et al[11]. The immune cell types vary in the number of gene markers defining each type with some cell types more extensively annotated than others, therefore providing more confidence in the calls of those. Sample-wise immune cell enrichment scores were calculated using raw read counts with Bioconductor R package *GSVA*[12]. The resulted enrichment scores of pre- and post-treatment samples in PL, LD and HD groups were compared separately using *limma*[13] by specifying a paired sample design. The immune cell types that met the following criteria were selected and displayed: 1. BH-adjusted *P-value*≤0.2 in HD; 2. No restriction on BH-adjusted *P-value* in LD; 3. If t-statistics in PL is the same direction as in HD, BH-adjusted *P-value* in PL must be >0.2; 4. If t-statistics in placebo is the opposite direction to HD, there is no restriction on BH-adjusted *P-value* in PL.

Co-clinical trial in two mouse cohorts of a Lynch syndrome model and assessment of whole genome expression in mice intestinal samples. A total of six 6-week-old males and females *Msh2*$^{LoxP/LoxP}$*;Villin-Cre*[14] mice were generously provided by Winfried Edelmann from Albert Einstein Medical College and were randomly assigned to treatment with naproxen at a dose of 30 mg/kg/day (N=3) or control (sesame oil, N=3). This dose is equivalent in a human (human equivalent dose, HED) of 70 kg of body weight to 175 mg q.d. and thus comparable to the LD naproxen level. Treatment and control were administered by oral gavage (daily oral administration via oral canula) over a period of 24 weeks. Animals were weighed every two weeks and a weight loss ≥20% was an indication for early euthanasia. At the date of termination of the experiment, mice were euthanized, the gut was dissected, flushed with phosphate

Reyes (2020)

buffered saline, and stored in RNA-Later at -80C. Animal experiments were carried out in compliance and approved by the Animal Care and Use Committee at MDACC.

Frozen samples preserved in RNALater were thawed on ice. RNA was extracted using Direct-zol RNA MiniPrep Plus kit (Zymo Research) according to the manufacturer's instructions. Column DNA digestion was performed using DNase I set (Zymo Research). RNA was eluted in nuclease free water and RNA concentration was measured using NanoDrop. Samples were stored in -80C until they were submitted for RNA-sequencing using Illumina HiSeq4000 instrument that generated 75-bp paired-end reads.

In regards to the bioinformatic analysis of mice samples, intestinal samples were sequenced at the MDACC SMF using an Illumina Hiseq4000 instrument that generated paired-end 75-base pair (bp) reads. Overall sequencing quality was evaluated using *FastQC*. Gene expression level was quantified using *rsem-calculate-expression* in *RSEM v1.2.312*[15], in which *STAR 2.5.2b*[5] was launched for aligning reads to NCBI GRCm38 mice genome reference. The gene level counts estimated according to GENCODE vM16 gene annotation of all samples were combined for downstream analyses.

Differential expression analysis between control and naproxen-treated mice samples was performed using Bioconductor R package *DESeq2*[7]. As in human mRNAseq analysis, PCA was performed. The significant genes with BH-adjusted p-value≤0.05 and absolute value of log2-Fold Change≥0.5 were highlighted and annotated with pathways of interest in a volcano plot.

Another cohort of $Msh2^{LoxP/LoxP}$;$Villin\text{-}Cre$[14] mice was used to perform a survival comparative analysis of naproxen, aspirin, and placebo. All mice were bred in a barrier facility according to IACUC guidelines and approved protocol by Weill Cornell Medical College and MD Anderson Cancer Center. Aspirin and naproxen were purchased from Spectrum Chemicals in powder form. Control mice were given pelleted plain RD Western Diet and treatment groups were given either aspirin 400 ppm (HED of 370 mg/day for

a human of 70 kg of weight), or naproxen 166 ppm (HED of 150 mg/day) impregnated pelleted RD Western Diets (Research Diets Inc) at the age of weaning that were identical, except for the presence of aspirin or naproxen. We monitored closely the amount of diet given to mice and also checked the weights of the mice during the study that did not differ between control, aspirin or naproxen arms. Overall, the naproxen dose used is closer to the LD treatment arm in the human clinical trial. Each treatment arm was populated with 30-40 mice in total keeping the same ratio of males and females. *Msh2*$^{LoxP/LoxP}$*;Villin-Cre*^{14} mice sporadically develop intestinal tumors over their life time sporadically. Accordingly, the health condition of the animals deteriorates over time. Mice were monitored throughout their life span for the following signs for determination of euthanasia: (1) weight loss, (2) poor coat quality, (3) hunched posture, and (4) pale limbs (anemia). Following IACUC guidelines, mice presenting with any of these signs were euthanized by CO2 inhalation. Subsequently, tissues were collected, and tumor burden (tumor multiplicity x weight) data were recorded. Histopathological evaluation was performed by an in-house pathologist and based on H&E staining of the tissues. All statistical analyses of survival were performed using GraphPad Prism 7.0. Survival data from the Kaplan-Meier (KM) survival curves were compared with the log-rank test and tumor burden plots were analyzed using Mann-Whitney non-parametric test. In all analyses, a value of P-value≤ 0.05 was considered statistically significant.

Derivation of patient derived organoids (PDO) for ex-vivo culture, treatment with Naproxen, whole transcriptomic analysis and qPCR validation analyses. A total of six PDOs lines were derived from normal colorectal mucosa of the rectosigmoid of six LS patients recruited during their routine screening colonoscopy to an IRB-approved biospecimen protocol (MDACC IRB# PA12-0327). These participants were independent from the 'Naproxen trial' and did not overlap. The process to establish organoid cultures has been previously described[16]. In brief, tissue pieces were washed three times with 1x cold PBS to remove contaminants such as hair and feces. Then, tissue pieces were incubated in a chelating buffer containing 1x PBS (without Ca^{+}/Mg^{+}) and 2mM EDTA at 4 ºC for 30 min on a rotating platform. Tissue pieces were then transferred to fresh cold 1xPBS and crypts were dissociated by vigorously shaking the

Reyes (2020)

tube ten times. Dissociated crypts were spun down at 300 g for 10 min at 4 ºC. Whole crypts were suspended in Matrigel and seeded in a pre-warmed 24 well culture plate as 50ul drop per well. Polymerized matrigel domes were bathed in 500 µl of cocktail medium containing conditioned media for Wnt3a, R-Spondin, and Noggin recombinant proteins, and other growth factors essential for expansion and maintenance of the organoids.

All of the PDO lines were treated with either DMSO (control), 500 uM, and 1 mM naproxen for 48 hours in 24 well plates. The dose of 500 uM was selected based on the assessment of naproxen levels in plasma in participants to the 'Naproxen trial'. HD naproxen participants had a plasma concentration of approximately 300 uM (**Table S8**, range 160 – 592 uM). However, since PDOs present cystic structures there were concerns for the drug reaching clinically significant concentrations in the inner epithelial layers. Therefore, two doses of 500 uM and the 1 mM were selected in order to achieve the targeted dose, to establish a dose-response pattern, and to assess cytoxicity of the agent on normal colorectal epithelium. We did not observe cytotoxicity even at a concentration of 1 mM using a standard MTT assay. For each treatment, PDOs from at least six wells were treated. At the end of the treatment, organoids were micrographed using brightfield microscope. Culture medium was carefully removed and organoids were collected by pipetting several times in cell recovery solution and incubated for 45 min on ice to remove Matrigel by de-polymerization. Treated organoids were centrifuged at 300 g for 5 min and pellet is washed once with 10 ml ice cold basal culture medium (DMEM:F12). Upon centrifugation, Trizol reagent was added directly to the pellets and total RNA was isolated using RNA isolation kit (Ambion. Life Technologies). Six lines were used for mRNAseq and 4 were used for validation of specific genes using qRT-PCR.

In regards to mRNAseq analyses, PDOs samples were sequenced at the MDACC Sequencing and MicroArray Facility (SMF) using an Illumina HiSeq4000 instrument that generated paired-end 75-base pair (bp) reads. Overall sequencing quality was evaluated using *FastQC*. Gene expression levels were

quantified using *rsem-calculate-expression* in *RSEM v1.2.212*[15], in which *Bowtie*[17] by default was launched for aligning reads to UCSC hg19 human genome reference. Gene level counts estimated according to GENCODE v19 gene annotation of all samples were combined for downstream analyses. Differential expression analysis between DMSO (control) and naproxen-treated PDO was performed using Bioconductor R package *DESeq2*[7] by specifying a paired sample design. As in human mRNAseq analysis, PCA was performed first. Significant genes with BH-adjusted *P-value*≤0.05 and an absolute value of log2-Fold Change≥0.5 were highlighted and annotated with pathways of interest in volcano plots. Then, mRNA expression of key genes identified in mRNAseq and human data was validated by qRT-PCR analysis following previously described method for validation purposes[18]. Briefly, total RNA was isolated using Trizol and 200 ng of total RNA was reverse transcribed into cDNA using High capacity reverse transcriptase kit. Authenticated PCR primers were used for the qRT-PCR assay of selected genes using Applied Biosystems's Fast SYBR green mastermix on Applied Biosystems Viia 7 Real-Time PCR System (Life Technologies Inc.) Data analysis was performed using Ct mean and expressed in $2^{-\triangle Ct}$. The data are expressed as means ±SD from three technical replicates and from four different organoid lines. Primer sequences are included in the **Supplementary Information** (**Table S3**).

Comparative Analyses of LD and HD Naproxen in Human to Organoid and Mice Data Sets. Despite selecting the appropriate dose to treat human organoids and mice cohorts based on HED, we assessed the global correlation of the naproxen effects at the transcriptome level among human LD, HD, organoid and mice samples before proceeding to select the most adequate model for validating the LD and HD human signatures. Common genes present in all datasets were selected after converting mice ensemble gene IDs to human gene symbols using Bioconductor R package *biomaRt*[19 20]. Read counts of common genes in all datasets were combined and VST normalized using *DESeq2*[7]. Patient-wise fold change of genes in human datasets were calculated by dividing normalized expression in post-treatment samples by normalized expression in pre-treatment samples, while patient-wise fold change of genes in the organoid dataset were

calculated by diving normalized expression in naproxen treated samples to normalized expression in DMSO treated samples. Due to the unpaired nature of the design of the mice experiment, fold change of genes were calculated by dividing the mean normalized expression of naproxen treated samples by the mean normalized expression of control samples. Significant human LD and HD genes with BH-adjusted *P-value*≤0.1 were further selected from the resulted gene by patient fold change matrix to calculate patient-wise Pearson correlations, which were clustered and displayed in an unsupervised hierarchical clustering heat map using Bioconductor R package *ComplexHeatmap*[8].

In order to compare the effects of LD and HD naproxen in human to organoid and mice, the number of significant genes with BH-adjusted *P-value*≤0.1 in LD and HD naproxen specimens with respect to organoid and mice were displayed separately in Venn diagrams.

Finally, based on the results of Venn diagram and correlation analysis, LD significant genes that were annotated in the volcano plot and also significant in the organoid data set, HD significant genes that were annotated in the volcano plot and also significant in the mice data set (both with BH-adjusted *P-value*s≤0.1) were selected for unsupervised clustering of organoid and mice samples, respectively. The expression patterns of these genes in the organoid and mice data sets were displayed in unsupervised hierarchical clustering heat maps using Bioconductor R package *ComplexHeatmap*[8].

References

1. Murphey LJ, Williams MK, Sanchez SC, et al. Quantification of the major urinary metabolite of PGE2 by a liquid chromatographic/mass spectrometric assay: determination of cyclooxygenase-specific PGE2 synthesis in healthy humans and those with lung cancer. *Anal Biochem* 2004;334(2):266-75. doi: 10.1016/j.ab.2004.08.019 [published Online First: 2004/10/21]

2. Gross ND, Boyle JO, Morrow JD, et al. Levels of Prostaglandin E Metabolite, the Major Urinary Metabolite of Prostaglandin E₂, Are Increased in Smokers. *Clin Cancer Res* 2005;11(16):6087-93. doi: 10.1158/1078-0432.Ccr-05-0733

3. Gallo P, Fabbrocino S, Vinci F, et al. Multi-residue determination of non-steroidal anti-inflammatory drug residues in animal serum and plasma by HPLC and photo-diode array detection. *J Chromatogr Sci* 2006;44(10):585-90. [published Online First: 2007/01/27]

4. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34(17):i884-i90. doi: 10.1093/bioinformatics/bty560 [published Online First: 2018/11/14]

5. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29(1):15-21. doi: 10.1093/bioinformatics/bts635 [published Online First: 2012/10/30]

6. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31(2):166-9. doi: 10.1093/bioinformatics/btu638 [published Online First: 2014/09/28]

7. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15(12):550. doi: 10.1186/s13059-014-0550-8 [published Online First: 2014/12/18]

8. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32(18):2847-9. doi: 10.1093/bioinformatics/btw313 [published Online First: 2016/05/22]

9. Korotkevich G, Sukhov V, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv* 2019:060012. doi: 10.1101/060012

10. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;21(11):1350-6. doi: 10.1038/nm.3967 [published Online First: 2015/10/13]

11. Nirmal AJ, Regan T, Shih BB, et al. Immune Cell Gene Signatures for Profiling the Microenvironment of Solid Tumors. *Cancer Immunol Res* 2018;6(11):1388-400. doi: 10.1158/2326-6066.CIR-18-0342 [published Online First: 2018/09/30]

12. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7. doi: 10.1186/1471-2105-14-7

13. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43(7):e47. doi: 10.1093/nar/gkv007 [published Online First: 2015/01/22]

14. Kucherlapati MH, Lee K, Nguyen AA, et al. An Msh2 conditional knockout mouse for studying intestinal cancer and testing anticancer agents. *Gastroenterology* 2010;138(3):993-1002 e1. doi: 10.1053/j.gastro.2009.11.009

15. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011;12:323. doi: 10.1186/1471-2105-12-323

16. Sato T, Vries RG, Snippert HJ, et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 2009;459(7244):262-5. doi: 10.1038/nature07935 nature07935 [pii] [published Online First: 2009/03/31]

17. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10(3):R25. doi: 10.1186/gb-2009-10-3-r25

18. Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov* 2015;5(1):43-51. doi: 10.1158/2159-8290.CD-14-0863 [published Online First: 2014/11/02]

19. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 2005;21(16):3439-40. doi: 10.1093/bioinformatics/bti525 [published Online First: 2005/08/06]

20. Durinck S, Spellman PT, Birney E, et al. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009;4(8):1184-91. doi: 10.1038/nprot.2009.97 [published Online First: 2009/07/21]