Gene Expression Profiling and Machine Learning To Characterize the Molecular Mechanisms of PI3K Inhibitor-Induced Gastrointestinal Side Effects

Fullerton[2], J. Grahnen[1], W. Proctor[2], L. Pham[1], M. Tseng[2], J. Fine[2], B. Goldberg[1], P. Lum[1] and S. Chandler[2]
[1]Auransa Therapeutics, Palo Alto, CA; and [2]Genentech, South San Francisco, CA.

Introduction:

Iatrogenic gastrointestinal disorders are among the most debilitating side effects of cancer therapy. In many cases, mucositis, colitis, diarrhea, and other drug-induced conditions limit tolerability of chemotherapy. Some targeted therapies, such as PI3K inhibitors, are also known to cause gastrointestinal (GI) side effects. Thus, understanding the molecular mechanisms of GI toxicity caused by chemotherapeutic and targeted agents may provide insights into key genes and pathways, highlighting potential targets suitable for co-therapies that can mitigate them. To this end, we have used computational methods to explore the connections between GI diseases such as Crohn's disease (CD), inflammatory bowel disease (IBD), GI effects caused by bacteria and viruses, and GI toxicities caused by drugs or compounds. In total, we have investigated over 60 publicly available datasets that included compounds such as irinotecan, wortmannin, methotrexate and multiple PI3K inhibitors. We also performed whole-transcriptome profiling of the human EpiIntestinal™ in vitro model treated acutely with diverse PI3K inhibitors at their respective plasma steady-state concentrations. The data and conclusions presented here outline a novel approach to identify non-obvious connections and pathways that could potentially inform on both mechanisms of toxicity as well as potential strategies for mitigation.

Materials and Methods:

PI3K Inhibitor Compound Set:

Isoform Selectivity

Compounds:	p110α	p110β	p110δ	p110γ	Classification:
A	+++	+	+++	++	Beta sparing
B	++	++	++	++	Pan-Inhibitor
C	+++	+	+	+	Alpha selective
D	+	+	+++	+	Delta selective
E	+++	+	+	+	Alpha selective
F	++	++	++	++	Pan-Inhibitor
G	++	+	++	++	Beta sparing

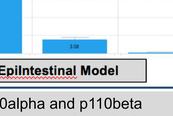
+ Low <------Potency------> +++ High

A set of 7 clinically relevant PI3K inhibitors were selected to comprise treatments with diverse isoform selectivity

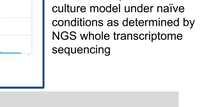
In-Vitro Model System:

- MatTek EpiIntestinal Culture

The in-vitro model system utilizes isolate primary human small intestinal cells to reconstitute a physiologically and morphologically relevant GI culture system containing GI epithelial cells, Goblet cells and fibroblasts. EpiIntestinal model has more relevant drug transporter expression, increased CYP450 metabolic activity, and barrier properties (lower tran-epithelial electrical resistance (TEER)) compared to Caco-2 cells.



Schematic Representation of MatTek tissue models



EpiIntestinal (SMI-100) exhibits appropriate epithelial structure and markers. Cytokeratin 19* Epithelial Cells (Red) and Villin (Green)

Images depicted here are from MAtTek supplied materials:

Experimental Design:

EpiIntestinal culture samples were allowed to acclimate for 24h after arrival. At time 0h on Day 1 samples were treated with either 1x or 10x the reported plasma steady state concentration of the respective PI3K inhibitor, which were previously established to be non-cytotoxic as indicated by MTT assessment and barrier function TEER analysis (data not shown). Samples were collected at 4 hrs and 24 hrs after compound treatment and processed to obtain RNA for whole transcriptome sequencing.

Whole Transcriptome Sequencing:

Total RNA was isolated from EpiIntestinal In-vitro tissue samples utilizing miRNA-Easy (Qiagen) and Trizol reagent according to manufactures protocol. RNA was purified by depletion of ribosomal RNA using Ribominus kit (Ambion) followed by generation of cDNA libraries using Ion Total RNA-Seq kit (Thermo Fisher) according to manufacture protocols. Samples were sequenced on a 540 Ion Torrent chip using an Ion Torrent S5 sequencer to obtain a read depth of >15 million reads per sample library. Subsequently samples were aligned to the human genome (HG38) and annotated to Ensembl 90 using Partek Flow software (Partek) to obtain read counts per gene. Read counts were then normalized across samples using Trimmed M-of the Mean to prepare for statistical analysis. Normalized gene read counts were transferred to Capella for further analysis.

Computational Approach

The SMarTR™ engine is comprised of an ensemble of machine learning, mathematical and statistical approaches.
It is highly automated and runs in the cloud. Analysis of gene expression data was conducted on the basis of mean log fold changes (logFCs) vs control. For the PI3K inhibitors, the control was taken to be DMSO treatment at the appropriate time point. For public data, the appropriate controls were chosen as indicated by the Gene Expression Omnibus (GEO) data description or any associated publication. Wherever applicable, subtypes of a given disease was identified via the engine.

Public data was obtained from GEO, from which >60 data sets with experimental systems suspected to mimic GI toxicity was downloaded. Some datasets were compound treatments and some were inflammatory diseases such as Crohn's disease. The data was appropriately transformed and logFCs vs control were computed as described above. To match the PI3K data, experiments were segregated by dose and time point when possible.

Similarity measures were derived via computation of correlations between experimental conductions across selected gene sets and subsequent clustering. To increase comparability, logFCs from individual experiments were Z-scored on a column basis, transforming all the data to a similar scale.

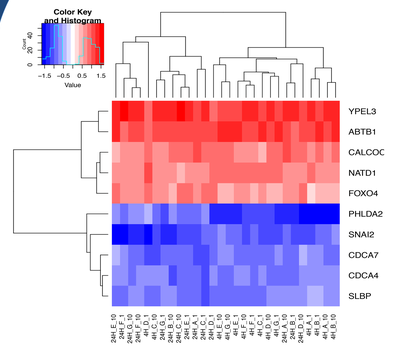
Results:

PIK3 isoform expression in Human Tissue and EpiIntenstinal:

Gene expression of PI3K isoforms: PIK3CA (p110alpha), PIK3CB (p110beta), PIK3CD (p110delta), PIK3CG (p110gamma) as reported by the NIH GTEX database and isoform expression in the MatTek EpiIntestinal culture model under naïve conditions as determined by NGS whole transcriptome sequencing

- EpiIntestinal culture samples express PI3KC isoforms p110alpha and p110beta
- EpiIntestinal cultures have lower p110delta and p110gamma expression, which are highly expressed in immune cells.

Core Genes of PI3K Inhibitors



Inhibitors Segregate by Isoform Selectivity

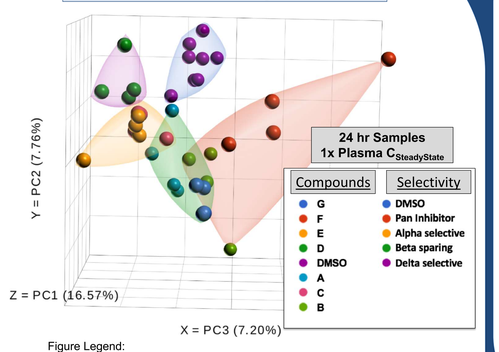


Figure Legend:
NGS analysis yielded transcriptome signatures that clearly segregate based on compound treatment and by isoform selectivity.

- In General, at clinically relevant exposures response across PI3K inhibitors in the EpiIntestinal model were similar.
- PCA analysis segregated molecules based on isoform selectivity

Similarity Matrix at the Transcriptome Level Identifies Similarities to *C. diff* 24 Hr Treatment

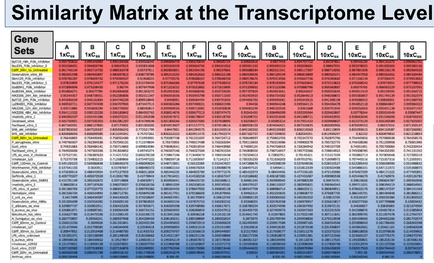


Figure Legend:
Similarity matrix between logFCs induced by PI3K inhibitors and preclinical models of GI toxicity

- Utilizing Capella's SMarTR™ engine, publicly curated information was matched to PI3K inhibitor transcriptional profiles to support high similarity in c. difficile associated pathway changes.
- The closest pre-clinical model for PI3K inhibitors is the treatment of Tox A of *C. difficile* in human ileocecal epithelial cells (GSE 29008).

Gene Expression Heatmap for Key Responsive Genes of PI3K Inhibitors

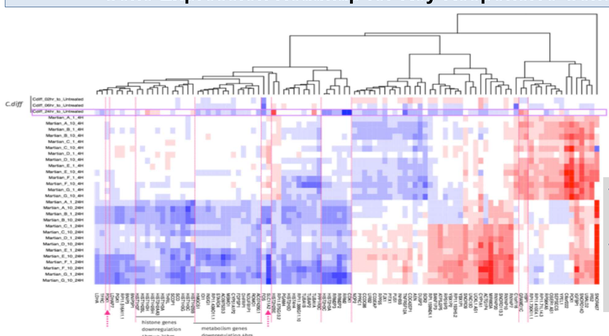


Figure Legend:
Heatmap of induced logFCs in genes which are significantly dysregulated by a plurality of PI3K inhibitors; PI3K inhibitors at various time points and doses (bottom rows) and C. diff preclinical model at various time points (top rows) are shown

- Histone and lipid metabolism genes are down-regulated upon treatment of cells with the inhibitors
- Transcriptional differences of 4hrs and 24hrs treatments are larger than differences between various PI3K inhibitor isoforms

Expression Pattern of Specific Autophagy Genes

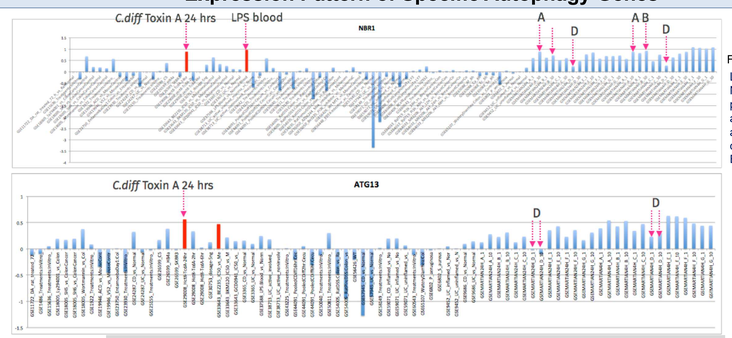


Figure Legend::
Log fold changes of NBR1 and ATG13 are plotted against publicly available datasets as well as treatments of compounds A-G in EpiIntestinal cells

- Key autophagy genes such as ATG13 and NBR1 are induced in a similar manner between *C.diff* Toxin A 24hr treatment and treatments with the various PI3K inhibitors
- Compound D does not elicit a robust up-regulation of the autophagy genes

Autophagy Genes Regulation Across PI3K Inhibitors and *C. diff*

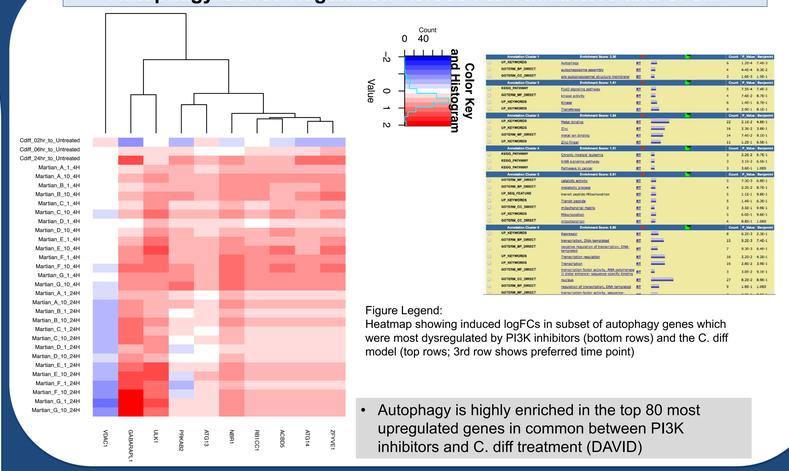


Figure Legend:
Heatmap showing induced logFCs in subset of autophagy genes which were most dysregulated by PI3K inhibitors (bottom rows) and the C. diff model (top rows; 3rd row shows preferred time point)

- Autophagy is highly enriched in the top 80 most upregulated genes in common between PI3K inhibitors and C. diff treatment (DAVID)

Genes and pathways involved

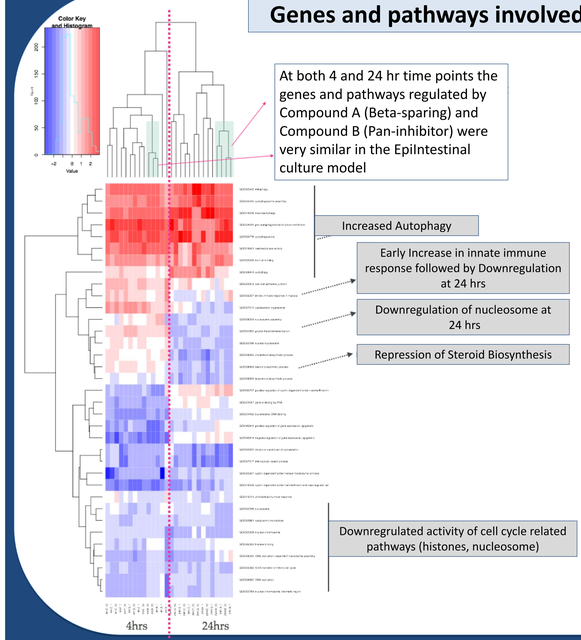


Figure Legend:
Heatmap of net pathway activations (here, a linear combination of logFCs in the gene set) for pathways which are significantly enriched for genes dysregulated by PI3K inhibitors. Rows mark GO pathway (gene set) IDs; columns mark PI3K inhibitor experiment. Experiments at the 4 hr time point to the left of the dotted line; those at the 24 hr time point on the right

CONCLUSIONS

1. We assessed transcriptional changes of a diverse panel of clinically relevant PI3K inhibitors in a reconstituted small intestinal epithelial model and utilized machine learning to map similarities of these gene changes to diverse and extensive publicly available data sets.

1. The closest match for PI3K inhibitor gene response was that of Tox A of *C. difficile*, thus supporting a non-obvious connection between an infectious disease etiology and PI3K inhibition related GI adverse events.

2. As part of these findings a common autophagy signature was observed across PI3K inhibitor molecules and Tox A response of *C. difficile*.

3. These findings present new hypothesis to test in regards to pathways resulting in PI3K inhibition mediated GI toxicities and potential mitigation strategies.