Dear Luca and David,

I’m reaching out in priority, as the original generators of the datasets:

## Selection of experiences/batches and replicates:

* In the provided `exp200921\_whole exp.tabular` file, it turned out that all drug samples related to the Drug Response batch are present, while some experiences in the more standard 200921 experience have been discarded. Report to replicates\_discarded\_200921.txt and `replicates\_preserved\_200921.txt attached for details (in total, 4 replicates times 5 compounds (Palboc, Olapar, LGK9743, Chloro, Bafilo)=20 replicates overall).
* If I have well understood your mail, should I completely discard out all replicates paired with `T25` pass in `exp281022` (4 replicates times 4 compounds= 16 replicates overall) + `p42` and `p43`!? Subsequently, what does this `T25` refer to? At least from the global Correlation Drug Heatmap matrix.
* Same remark for the `exp130921 in vivo` experiment. Not included so far to compute drug similarities in the resulting network.
* Besides, can you confirm that I have the most comprehensive barcode counts profiles (namely all replicates AND all barcode IDs), for all experiences? Acknowledging that with R programming language, I’m not suffering from the intrinsic memory-limitations of Excel. I emphasised again on this point, as the provided original `.TABULAR` file was more comprehensive with respect to originally provided CSV or Excel files on Google Drive + there’s a strong discrepancy in terms of surveyed barcode IDs from one batch to another, going from only 93,132 unique barcode IDs in batch exp07022, up to 1,067,031 unique barcode IDs in batch `exp200921\_dose response osim`??
  + Even better, can you check that the reported number of rows (unique barcode IDs) and columns (replicates) provided as an attached `global\_summaries\_rows\_by\_cols.xlsx` file truly reflects your most original database?
* Finally, other spotted inconsistencies between provided general metadata `Table of compounds.xlsx` file and actual barcode replicate IDs are listed under <https://github.com/VeraPancaldiLab/DNABarcode-DrugFingerprint/issues>/2 .

## Metadata inconsistencies

* I do not know exactly what to do with paired drug effect Osimertinb+sorafenib: 0,1uM+5uM, as the only experience trying to evaluate the synergistic effect of two drugs.
* I spotted some replicates’ names associated with the `PEM` prefix in ` exp281022\_time course` experiment. Can you confirm that it stands for `Compound: Pemetrexed`? (Indeed, for the other experiences, it’s usually prefix `Pemetr` which has been used!!).
* Finally, other inconsistencies are reported under <https://github.com/VeraPancaldiLab/DNABarcode-DrugFingerprint/issues/1> .