RESEARCH ARTICLE

Figures and figure supplements

A statistical framework for assessing pharmacological responses and biomarkers using uncertainty estimates

Dennis Wang *et al*

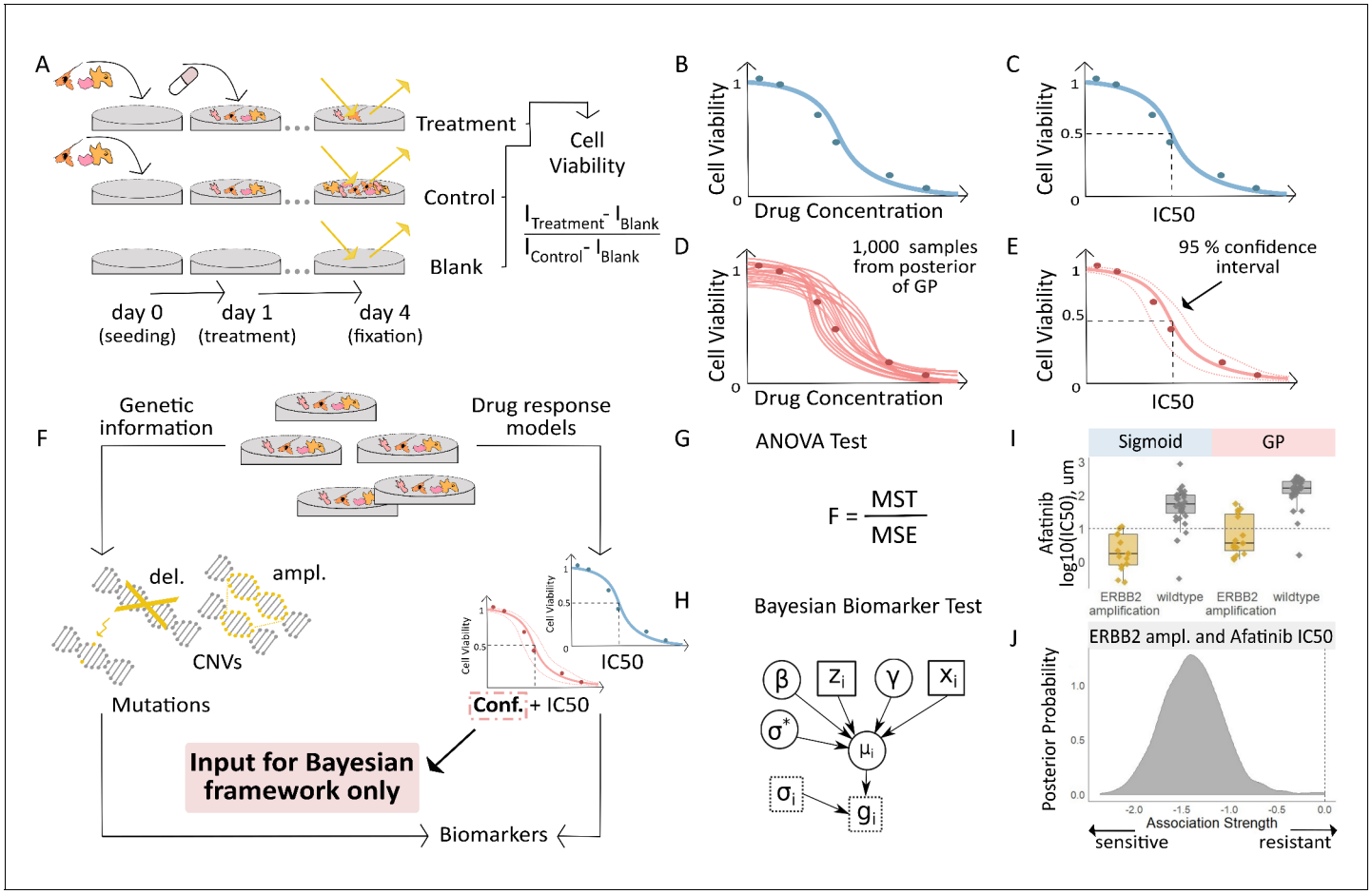


Figure 1. Workflow for fitting of Gaussian Process models to dose-response curves and estimating their uncertainty. (**A**) Large-scale drug screens test cell lines with different drugs and at different doses are used to obtain dose-response data. (**B**) Typically, for each drug tested in a cell line, the sigmoid model is fit to the drug-response data and (**C**) the overall measures of response (IC_{50}, AUC, etc.) are extracted. (**D**) For each drug tested in a cell line, we fit a GP model to the dose-response data. The GP allows us to sample from a distribution of possible dose-response curves, obtaining a measure of uncertainty. (**E**) From these curves, we can extract overall measures of response, such as IC_{50}, and importantly, their 95% confidence intervals. (**F**) Mutation markers for each cell line can be determined based on presence/absence of single nucleotide polymorphisms (SNPs) in key genes. Both the drug-response estimates and the mutation markers are used to compute (**G**) the F-statistic for ANOVA, and (**H**) Bayesian test for biomarker association. The drug-response summary measure g_i for cell line i is modelled via a cell line- specific mean μ_i and standard error σ_i. The mean is defined as a linear effect β of the biomarker status z_i and a further effect γ from any remaining covariates x_i, such as tissue type. The parameter σ^* is the standard deviation of μ_i. (**I**) Boxplots illustrate the differences in the estimated mean IC_{50} of *ERBB2* amplified and non-amplified breast cancer cell lines treated with afatinib. An ANOVA test was used to test this difference in means but did not consider uncertainty in each IC_{50} estimate. (**J**) We estimated posterior distributions of gene association using the Bayesian model, that is the effect of a genetic mutation on the IC_{50} measurement of drug response. Distributions centred on zero indicate no effect whilst distributions on either side of zero indicate positive or negative effects of mutations on drug response.

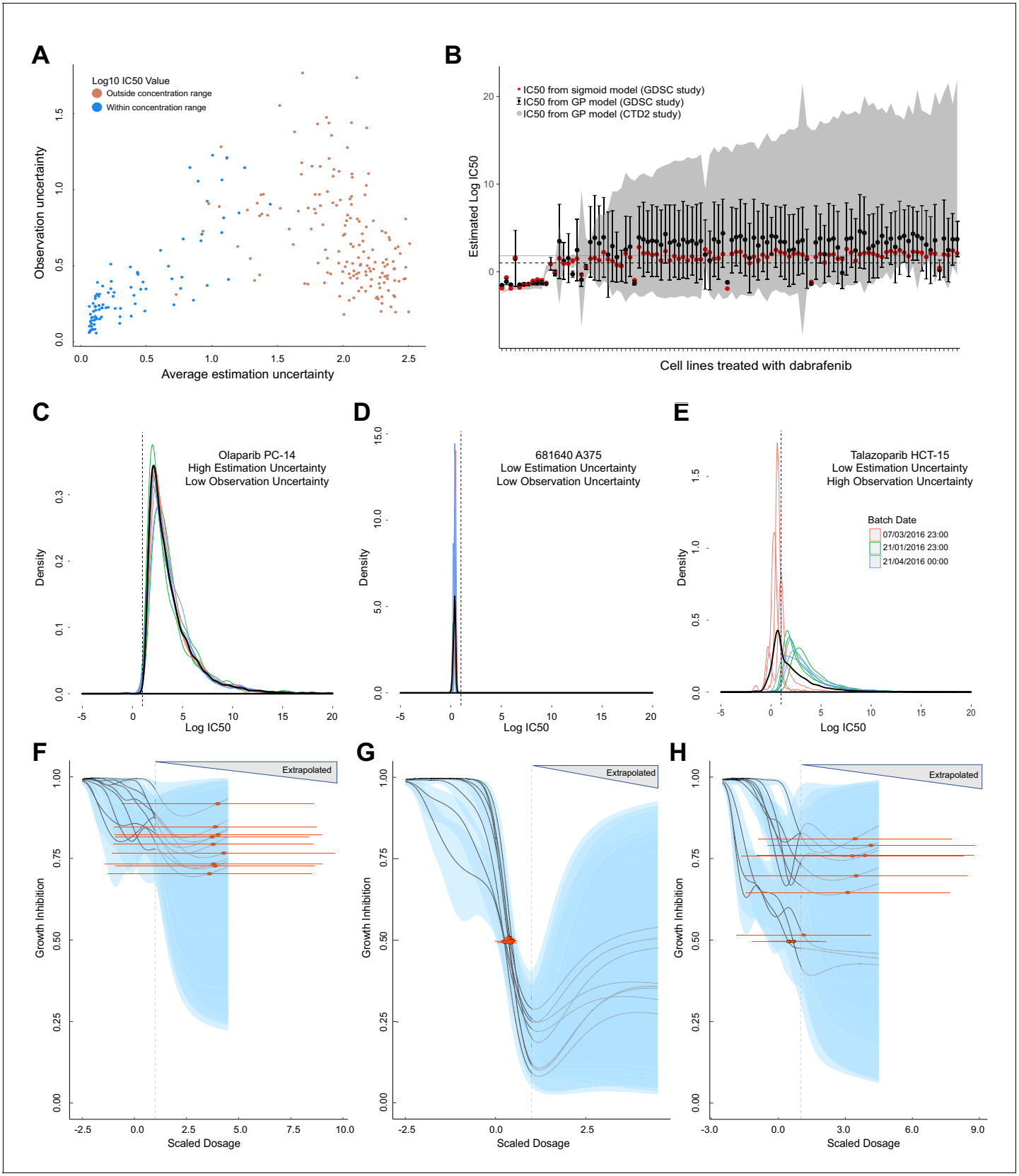


Figure 2. Comparison of GP estimates of uncertainty to replicate drug screening experiments. (**A**) Comparison between observational uncertainty (standard deviation over replicates of log10(IC_{50}) mean estimates) and estimation uncertainty (average over replicates of log10(IC_{50}) standard deviation)

Figure 2 continued on next page

Figure 2 continued

from each replication experiment. The colour of the points indicates whether the log10(IC_{50}) mean estimates were within or outside the maximum concentration range for each assay. (**B**) Mean IC_{50} and the estimation uncertainty from the GPs for a BRAF inhibitor (dabrafenib) tested in each cell line in two independent studies (GDSC and CTD2). Estimation uncertainty (error bars and grey shading) were larger beyond the max concentration in both GDSC (dashed line) and CTD2 (grey line). The point estimates of the IC_{50}s from the GPs (black dots) were also comparable to the published IC_{50}s (red dots). (**C-E**) Three sets of replicate experiments, representing different amounts of estimation and observation uncertainty. Each density represents the distribution of IC_{50} values from the Gaussian process samples from each replicate experiment. The colours represent different experimental batches. Narrow distributions demonstrate low estimation uncertainty and overlapping distributions demonstrate low observation uncertainty. The thick black line represents the density obtained by pooling samples from all replicates and the dashed line shows the maximal dosage tested. GP-curve fits corresponding to the three sets of replicate experiments showing IC_{50} estimates with (**F**) high uncertainty, (**G**) low uncertainty, and (**H**) mix of uncertainties depending on whether estimates are made within or beyond the max concentration. The blue areas represent the 95% confidence interval in the curve fits and extrapolated GP curves (light grey lines) are displayed up to five times the maximum concentration, where the uncertainty will be extremely high.

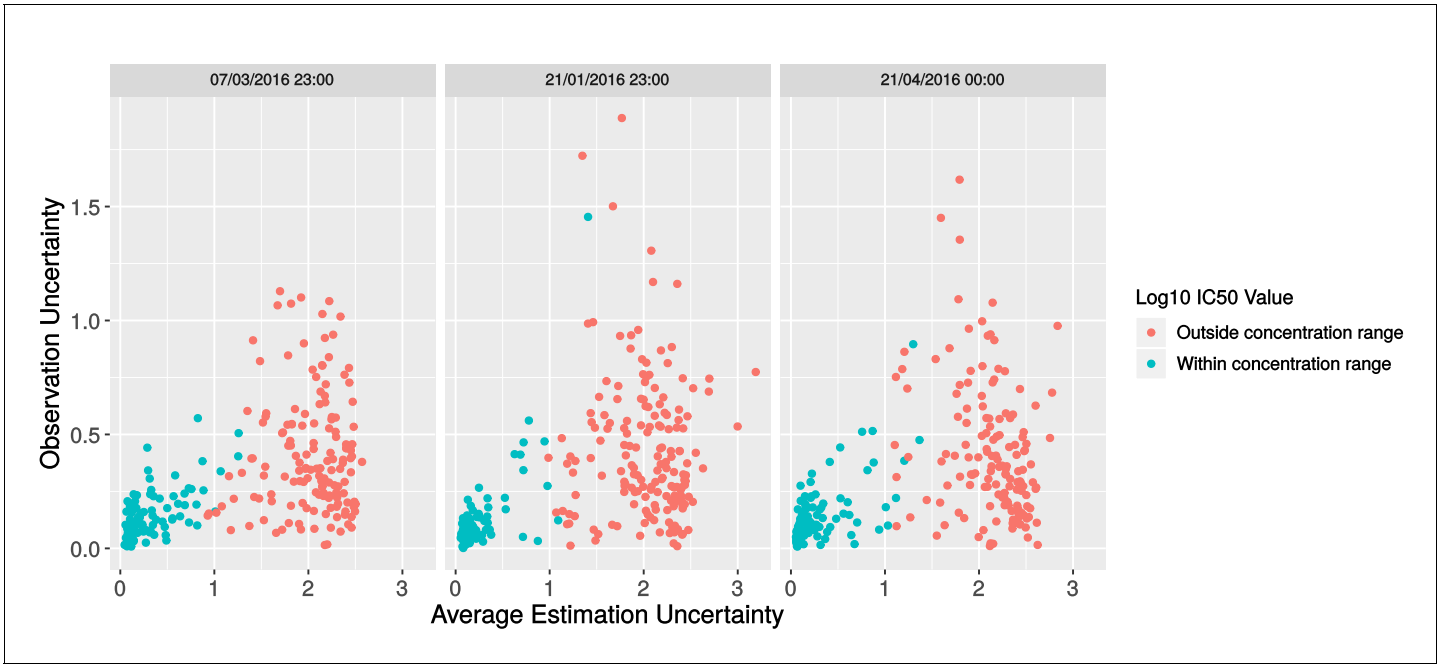


Figure 2—figure supplement 1. Investigation of batch effects in the replicate data. Scatterplot of observation uncertainty against average estimation uncertainty, split by experimental batch.

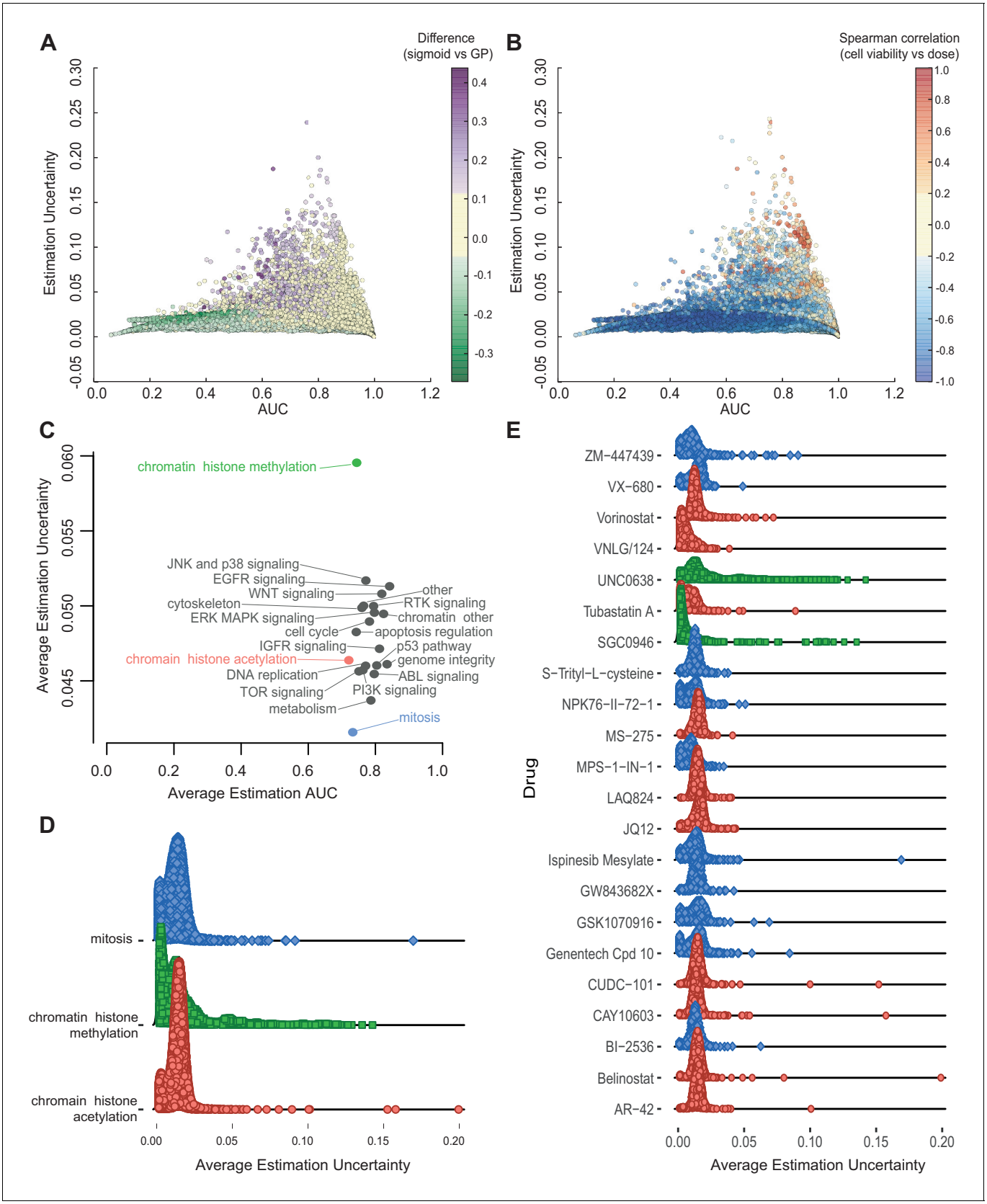


Figure 3. Relationship between AUC and uncertainties estimated from GPs across all experiments. (**A**) Coloured by difference between the AUC estimated by sigmoid vs GP fits. (**B**) Coloured by Spearman correlation between cell viability and dose concentration in the raw data. Poorer

Figure 3 continued on next page

Figure 3 continued

experiments (orange-red) result in greater uncertainty and positively correlated with cell viability increasing with higher dose. (**C**) Average uncertainty and AUC for experiments with uncertain fits (estimation uncertainty >0.03) with drugs grouped by their target pathway. (**D**) Distribution of estimation uncertainty for all drugs targeting chromatin histone methylation, chromatin histone acetylation, and mitosis and (**E**) for individual drugs.

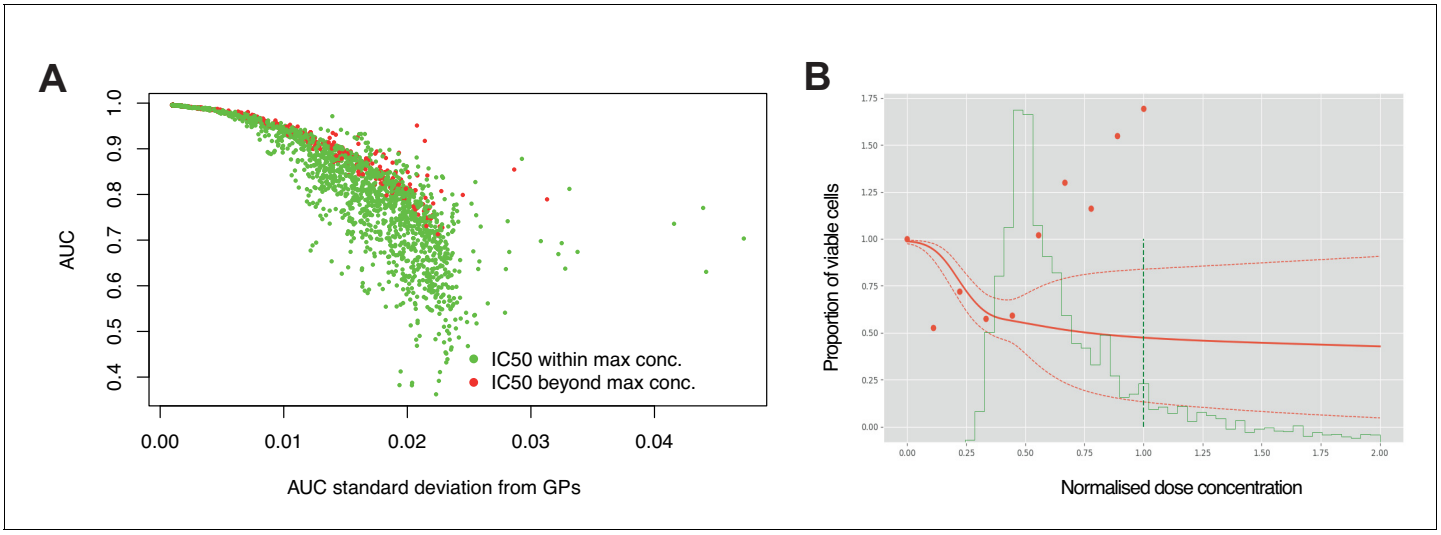


Figure 3—figure supplement 1. High estimation uncertainty independent of concentration range. (**A**) Estimation uncertainty for AUC (standard deviation) across replicate experiments were robust to whether the IC_{50} was within or beyond the maximum concentration tested. (**B**) Dose- response of a single experiment where there was high estimation uncertainty (dotted red line) despite the fitted curve (red line) crossing 50% viable cells before reaching the max dose concentration (dotted green line). The probability distribution of IC_{50} is estimated across dose concentrations (solid green line), however, raw data points (red dots) show increased cell viability with increased dose.

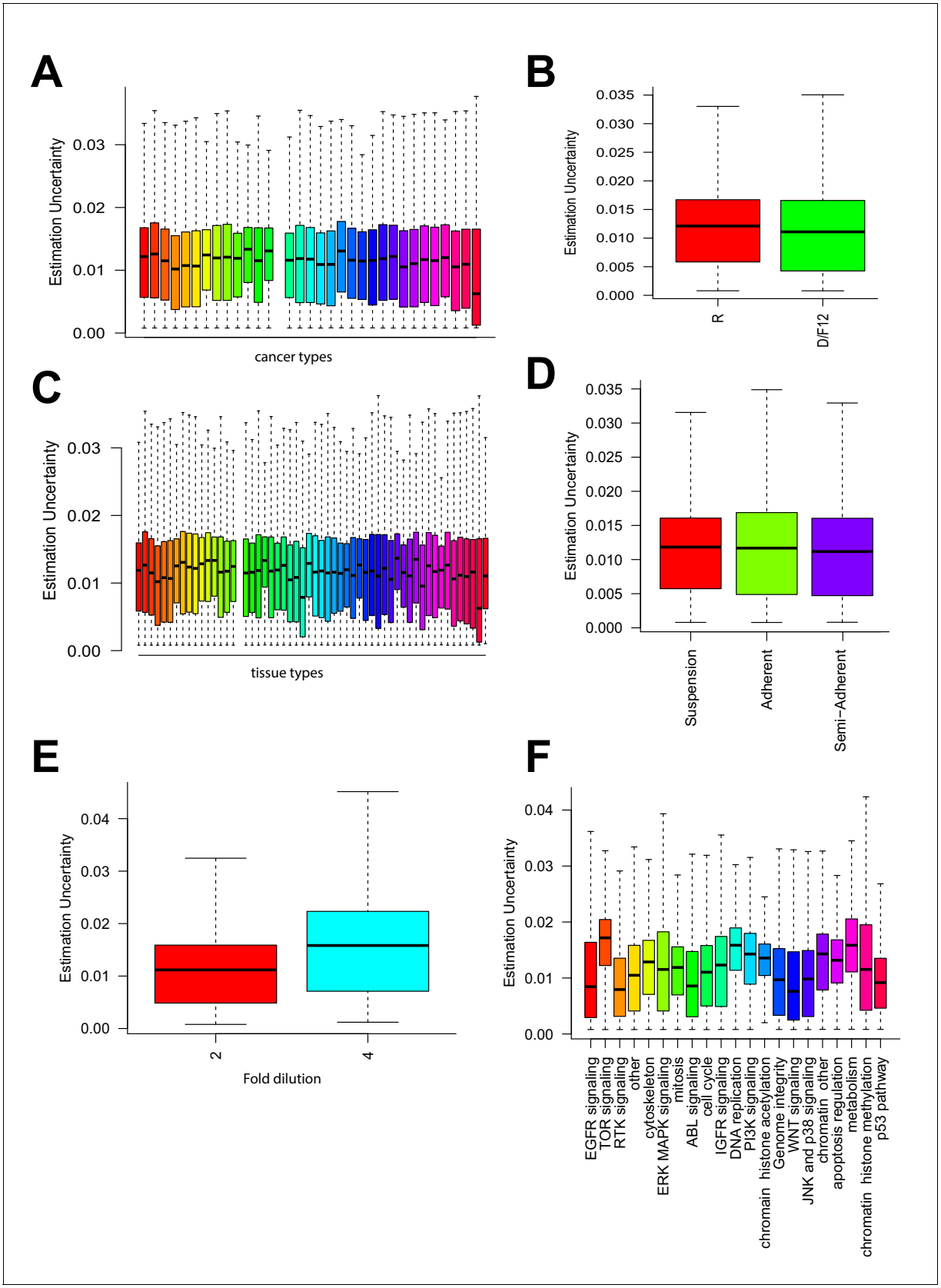

A
Estimation Uncertainty
0.00
0.01
0.02
0.03
cancer types
B
Estimation Uncertainty
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
R
D/F12
C
Estimation Uncertainty
0.00
0.01
0.02
0.03
tissue types
D
Estimation Uncertainty
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
Suspension
Adherent
Semi−Adherent
E
Estimation Uncertainty
0.00
0.01
0.02
0.03
0.04
2
4
Fold dilution
F
Estimation Uncertainty
0.00
0.01
0.02
0.03
0.04
EGFR signaling
TOR signaling
RTK signaling
other
cytoskeleton
ERK MAPK signaling
mitosis
ABL signaling
cell cycle
IGFR signaling
DNA replication
PI3K signaling
chromain histone acetylation
Genome integrity
WNT signaling
JNK and p38 signaling
chromatin other
apoptosis regulation
metabolism
chromatin histone methylation
p53 pathway

Figure 3—figure supplement 2. Relationship between uncertainties and other experimental factors. Estimation uncertainty grouped by (**A**) cancer type of cell lines, (**B**) growth media used, (**C**) tissue of origin, (**D**) growth condition of the cells, (**E**) dilution factor for each dose tested of a drug, and (**F**) target pathways of drugs.

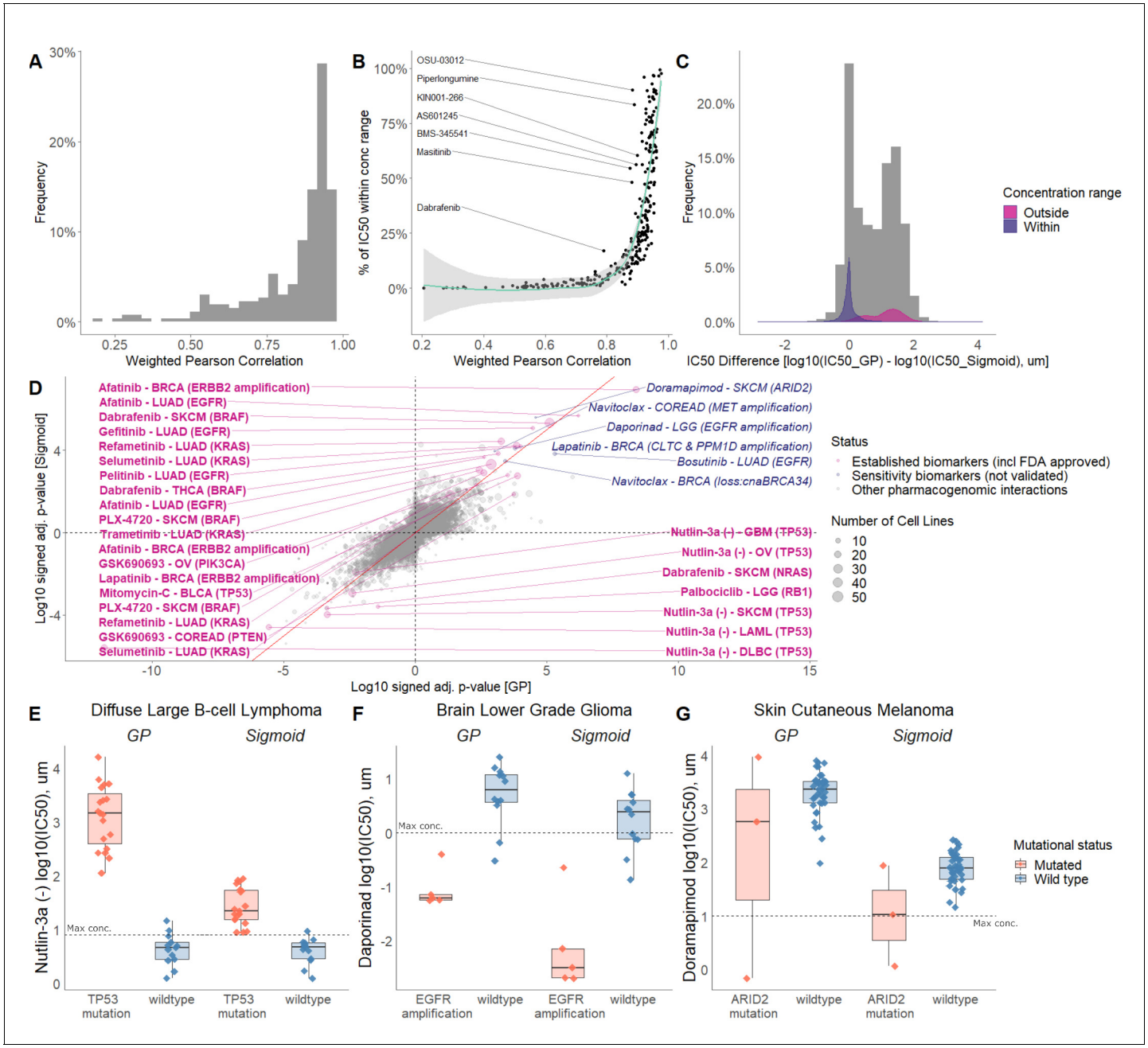


Figure 4. Comparison of sigmoid and GP-curve fitting. (**A**) Weighted Pearson correlation of each drug within cancer types. (**B**) Comparing the concordance of sigmoid and GP-curve fitting when stratifying for percentage of cell lines with IC_{50} value lower than maximum concentration. (**C**) IC_{50} value difference between GP and sigmoid curves. Grey histogram represents frequency distribution of the IC_{50} value difference between GP and sigmoid curves without stratification by within/outside the concentration range. (**D**) Drug-response biomarker comparison based on both curve fittings (sigmoid vs GP). The Benjamini-Hochberg adjusted p-values are in log10 scale and signed based on the direction of the effect size (Cohen's d). Additional biomarker examples for (**E**) diffuse large B-cell lymphoma (DLBCL) treated with nutlin-3a (MDM2 inhibitor) and stratified by *TP53* mutants; (**F**) Low grade glioma (LGG) treated with daporinad (NAMPT inhibitor) and stratified by *EGFR* amplification; (**G**) Skin cutaneous melanoma (SKCM) treated with doramapimod (p38 and JNK2 inhibitor) and stratified with *ARID2* mutations.

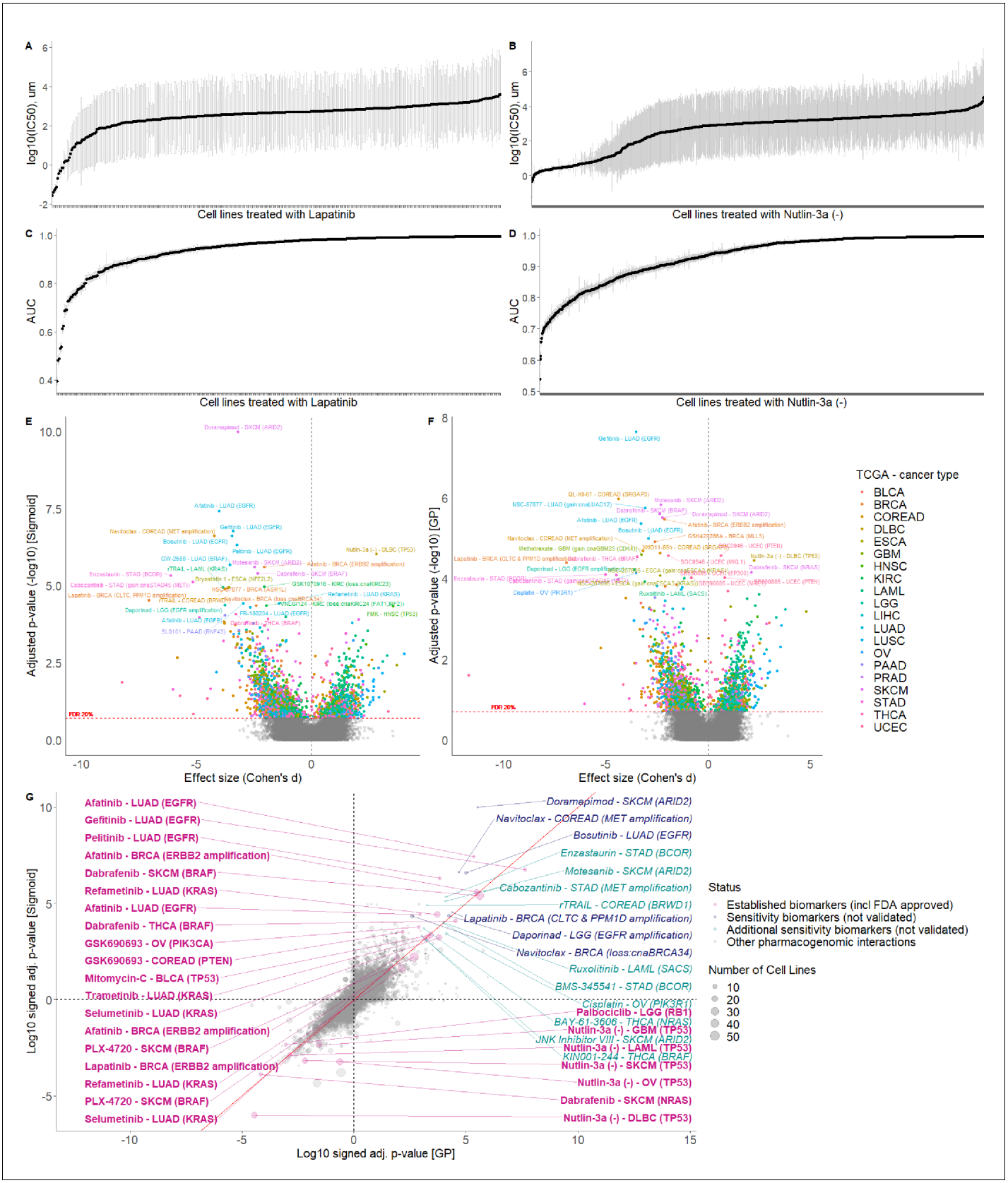


Figure 4—figure supplement 1. Comparison of GP and Sigmoid curve fitting using AUCs. GP fitted log10(IC50)s with standard error (grey error bars) of the cell lines treated with (**A**) Lapatinib and (**B**) Nutlin-3a. GP fitted AUCs with standard error (grey error bars) of the cell lines treated with (**C**)

Figure 4—figure supplement 1 continued on next page

Figure 4—figure supplement 1 continued

Lapatinib and (**D**) Nutlin-3a. Volcano plot of drug-response biomarker associations based on (**E**) sigmoid and (**F**) GP-curve fitting using AUC as a measure. (**G**) Drug-response biomarker comparison based on both curve fittings.

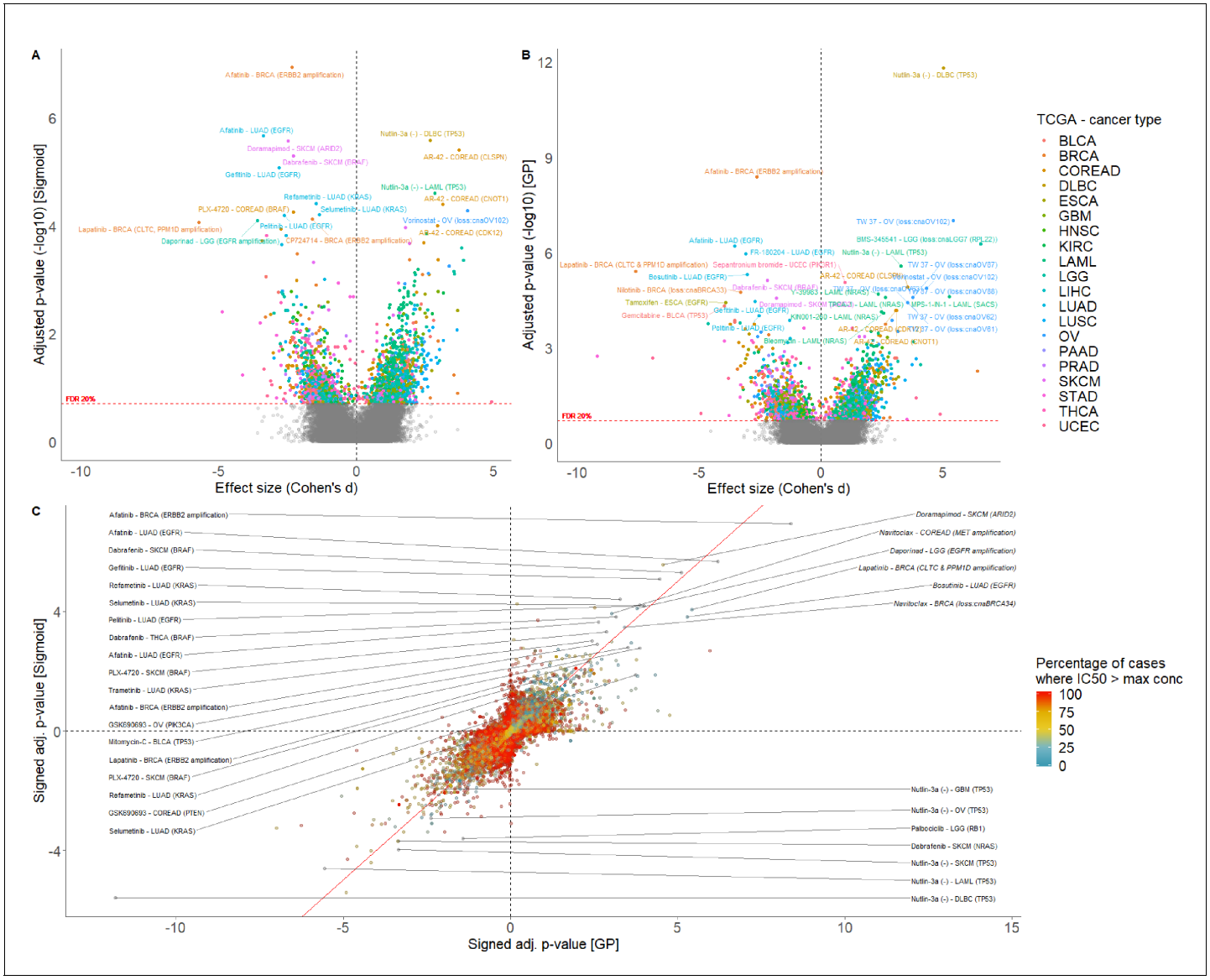


Figure 4—figure supplement 2. Comparison of sigmoid and GP-curve fitting using IC_{50}s. Volcano plot of drug-response biomarker associations based on (**A**) sigmoid and (**B**) GP-curve fitting. (**C**) Drug-response biomarker comparison based on both curve fittings, and colour coding percentage of drug-response data observed within- concentration range.

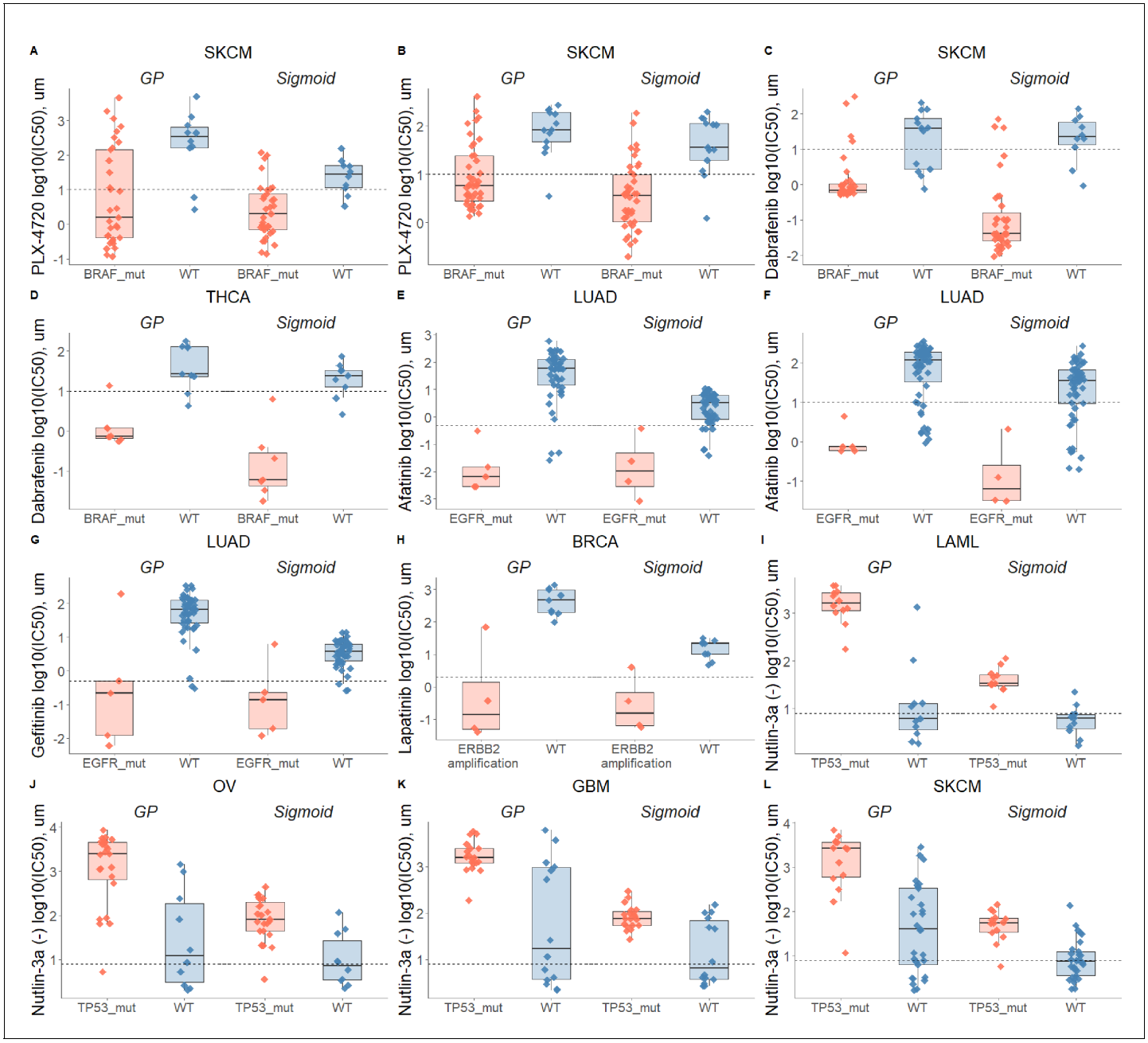


Figure 4—figure supplement 3. Drug-response biomarker comparison based on both curve fittings. Dashed line depicts maximum concentration. (**A**) Skin cutaneous melanoma (SKCM) treated with PLX-4720 (BRAF inhibitor) and stratified with *BRAF* mutations; (**B**) Skin cutaneous melanoma (SKCM) treated with PLX-4720 (BRAF inhibitor) and stratified with *BRAF* mutations - replicate; (**C**) Skin cutaneous melanoma (SKCM) treated with Dabrafenib (BRAF inhibitor) and stratified with *BRAF* mutations; (**D**) Thyroid carcinoma (THCA) treated with Dabrafenib (BRAF inhibitor) and stratified with *BRAF* mutations; (**E**) Lung adenocarcinoma (LUAD) treated with Afatinib (ERBB2, EGFR inhibitor) and stratified with *EGFR* mutations; (**F**) Lung adenocarcinoma (LUAD) treated with Afatinib (ERBB2, EGFR inhibitor) and stratified with *EGFR* mutations - replicate; (**G**) Lung adenocarcinoma (LUAD) treated with Gefitinib (EGFR inhibitor) and stratified with *EGFR* mutations; (**H**) Breast invasive carcinoma (BRCA) treated with Lapatinib (ERBB2, EGFR inhibitor) and stratified with *ERBB2* amplifications; (**I**) Acute Myeloid Leukemia (LAML) treated with nutlin-3a (MDM2 inhibitor) and stratified by *TP53* mutants; (**J**) Ovarian serous cystadenocarcinoma (OV) treated with nutlin-3a (MDM2 inhibitor) and stratified by *TP53* mutants; (**K**) Glioblastoma multiforme (GBM) treated with nutlin-3a (MDM2 inhibitor) and stratified by *TP53* mutants; (**L**) Skin cutaneous melanoma (SKCM) treated with nutlin-3a (MDM2 inhibitor) and stratified by *TP53* mutants.

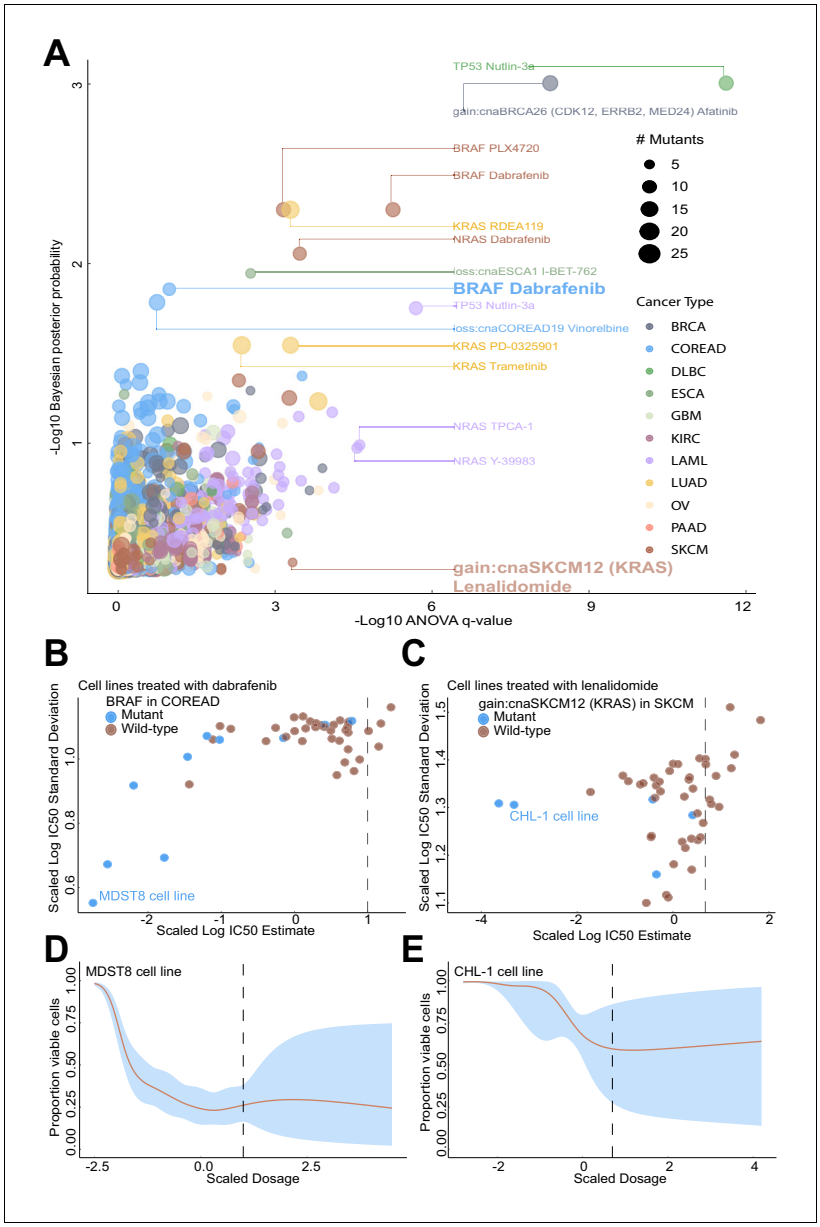


Figure 5. Comparison of Bayesian testing and ANOVA using the GP IC_{50} estimates. (**A**) Scatterplot of biomarker associations with IC_{50} drug response. The y-axis shows the negative log10 transformed posterior probability of a sign change in the effect under the Bayesian testing model, whilst the x-axis shows the negative log10 of the q-value from ANOVA testing. The size of the circles is proportional to the number of mutants or copy number variations in the given type of cancer cell line. (**B**) GP estimates for the mean and standard deviation of the log (IC_{50}) from colorectal cell lines tested with BRAF inhibitor dabrafenib, which showed significant association with *BRAF* mutation in the Bayesian test. (**C**) Estimated IC_{50} and its uncertainty for skin cutaneous melanoma cell lines tested with the immunomodulatory drug lenalidomide, which showed significant association with *KRAS* copy number alteration in the ANOVA test. Black vertical lines show the location of the maximum experimental drug dosage. Dose-response curve of the (**D**) MDST8 colorectal cancer cell lines with *BRAF* mutation treated with dabrafenib. The black dotted line represents the maximum concentration of the drug used to treat the cell lines. The blue area represents the 95% confidence intervals in the dose-response fits. (**E**) Similar to (**D**) but for CHL-1 skin cutaneous melanoma cell lines with *KRAS* copy number alteration treated with lenalidomide.

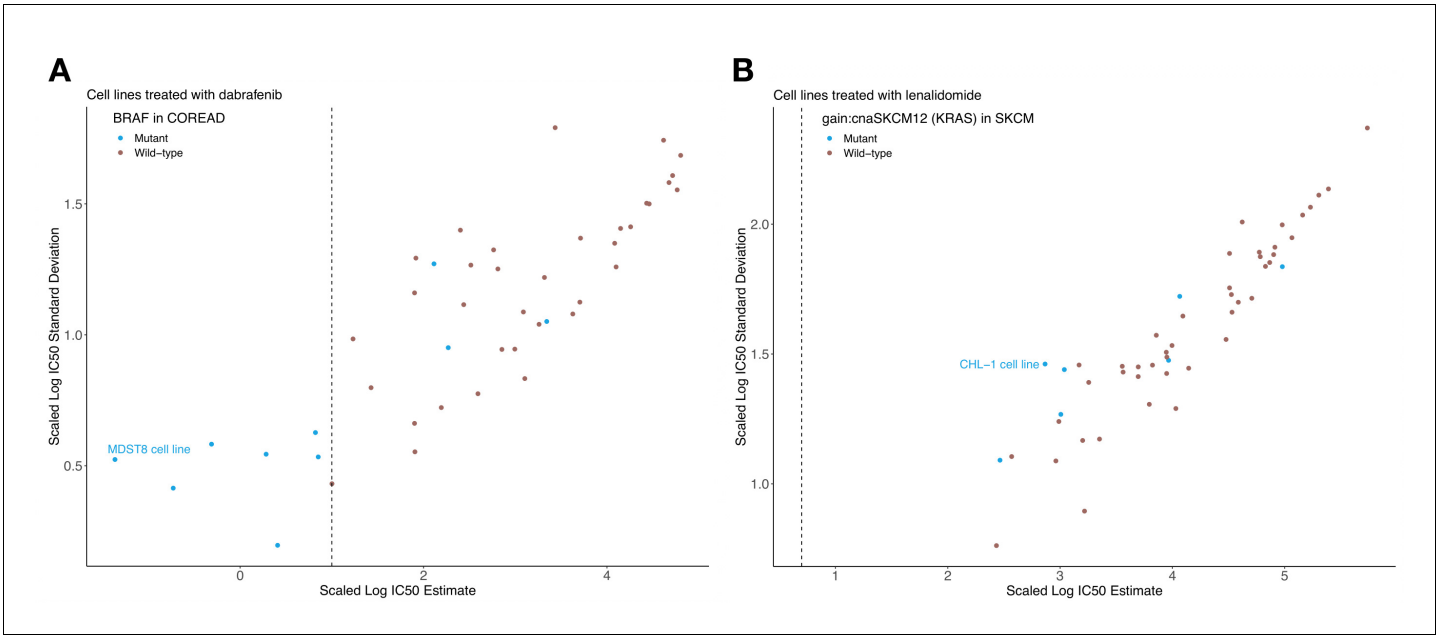


Figure 5—figure supplement 1. Sigmoid curve fitting uncertainty. Estimated log10(IC50) and uncertainties based on a bootstrap sampling method, for (**A**) colorectal cell lines tested with dabrafenib and (**B**) skin cutaneous melanoma cell lines tested with lenalidomide. Black vertical line represents the maximum experimental drug dosage.

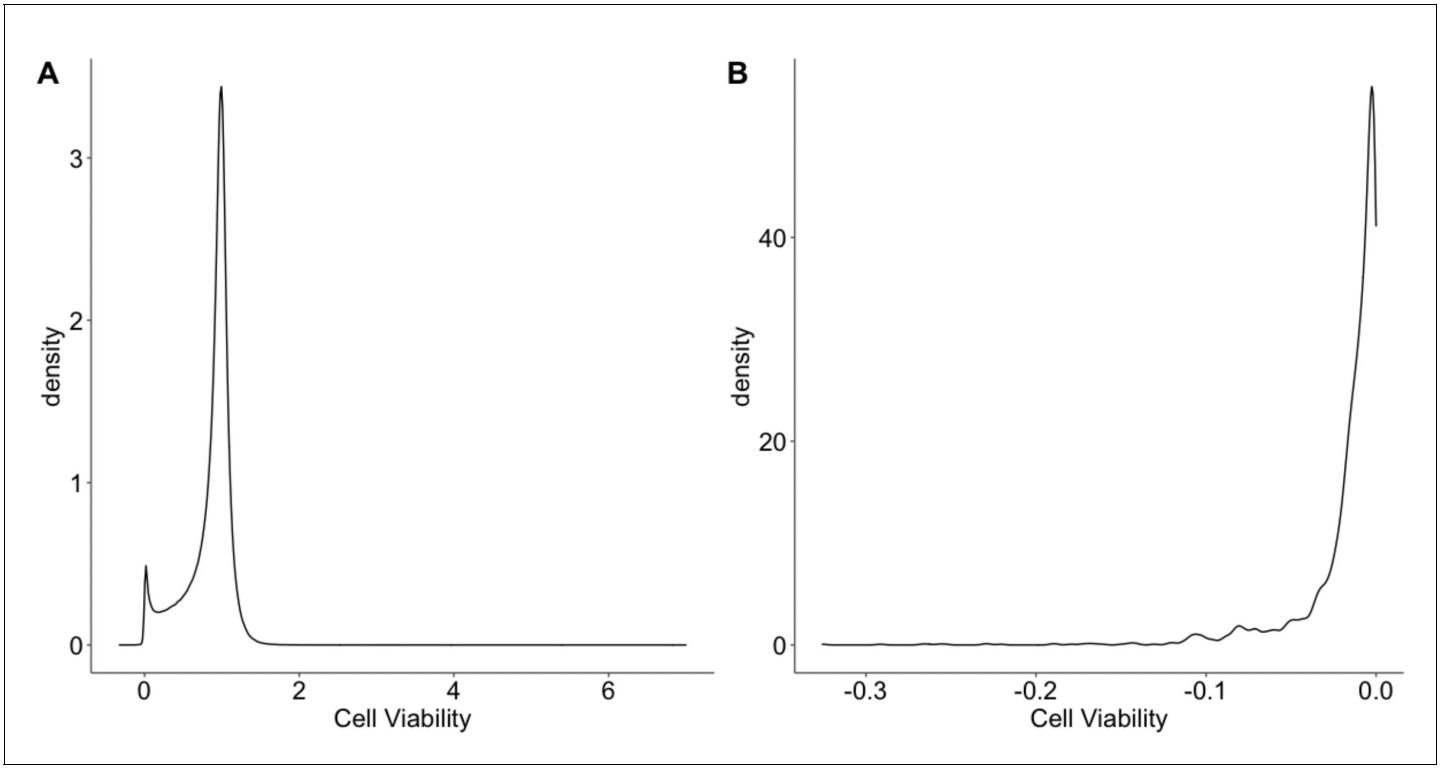


Figure 5—figure supplement 2. An overview of the cell viability values. (A) The distribution of the cell viability values. (B) The distribution of negative cell viability values.