

Figure 2 | Predictive modelling of pharmacological sensitivity using CCLE d, Specificity and sensitivity (receiver operating characteristic curves) of crossgenomic data. a, b, Drug responses for panobinostat (green) and PLX4720 validated categorical models predicting the response to a MEK inhibitor, PD-(orange/purple) represented by the high-concentration effect level (A_{max}) and 0325901 (activity area). Mean true positive rate and standard deviation (n = 5) transitional concentration (EC₅₀) for a sigmoidal fit to the response curve are shown when models are built using all lines (global categorical model, in blue and orange), or within only melanoma lines (green). e, Activity area values (b). c, Elastic net regression modelling of genomic features that predict sensitivity to PD-0325901. The bottom curve indicates drug response, for panobinostat between cell lines derived from haematopoietic (n = 61) and measured as the area over the dose-response curve (activity area), for each cell solid tumours (n = 387). The middle bar, median; box, inter-quartile range; line. The central heat map shows the CCLE features in the model (continuous bars extend to 1.5× the inter-quartile range. f, Distribution of activity area z-score for expression and copy number, dark red for discrete mutation calls), values for AEW541 relative to IGF1 mRNA expression. Orange dots, multiple across all cell lines (x axis). Bar plot (left): weight of the top predictive features myeloma cell lines (n = 14); blue dots, cell lines from other tumour types

(n = 434). Box-and-whisker plots show the activity area or mRNA expression

distributions relative to each cell line type (line, median; box, inter-quartile

range), with bars extending to 1.5× the inter-quartile range.

for sensitivity (bottom) or insensitivity (top). Parentheses indicate features

nnMS, non-neutral missense mutation (Supplementary Methods).

present in >80% of models after bootstrapping. LOF, loss of function mutation;