Supplementary Information

Comprehensive Characterization of Cytochrome P450 Isozyme Selectivity across Chemical Libraries

Henrike Veith[a], Noel Southall[a] Ruili Huang[a], Tim James[b], Darren Fayne[b], Natalia Artemenko[b], Min Shen[a], James Inglese[a], Christopher P. Austin[a], David G. Lloyd[b] and Douglas S. Auld[a]*

[a]NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD 20892-3370, USA

[b]Molecular Design Group, School of Biochemistry and Immunology, Trinity College Dublin, Dublin 2, Ireland

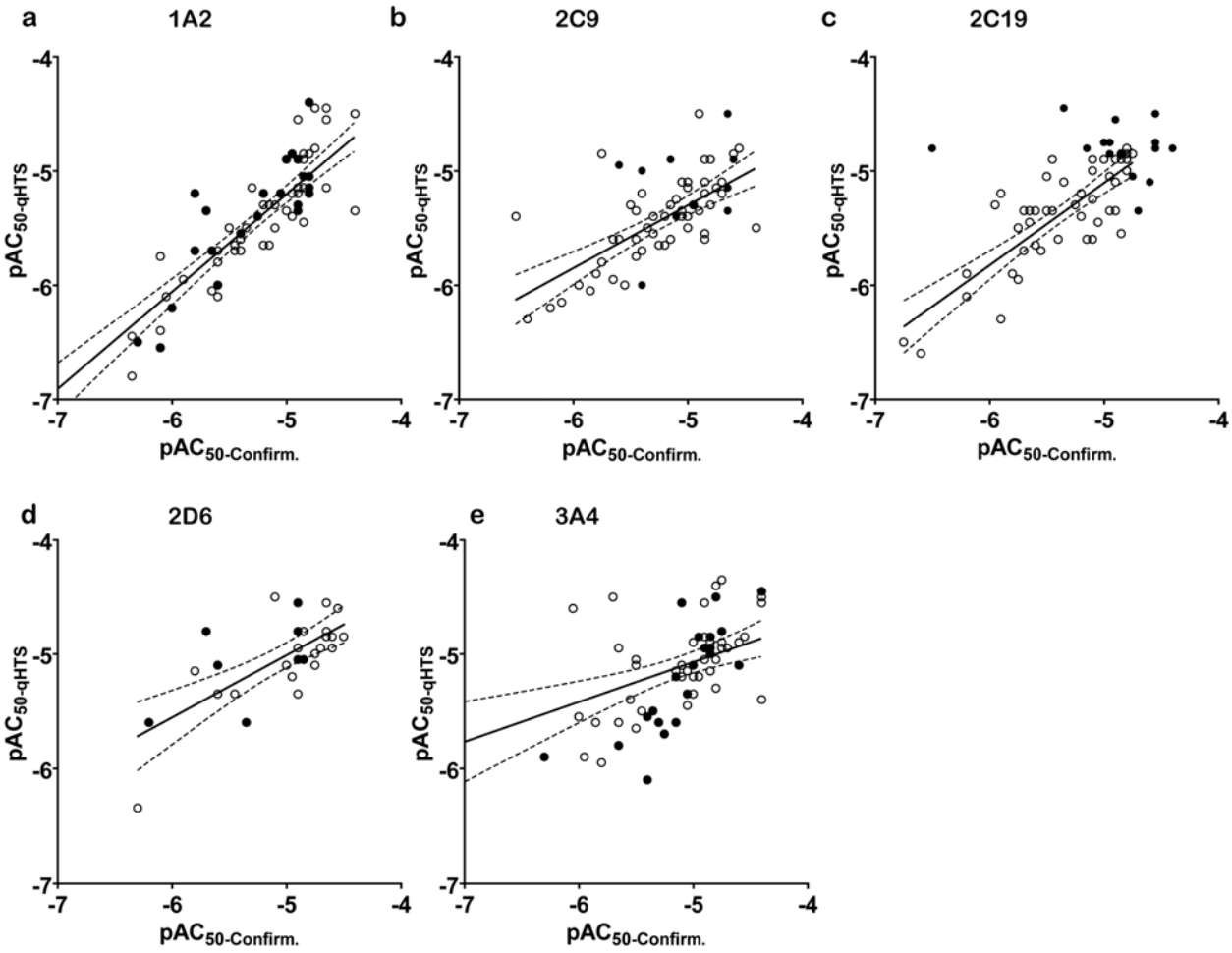
*Corresponding author

Running title: qHTS analysis of Cytochrome P450 isozyme selectivity

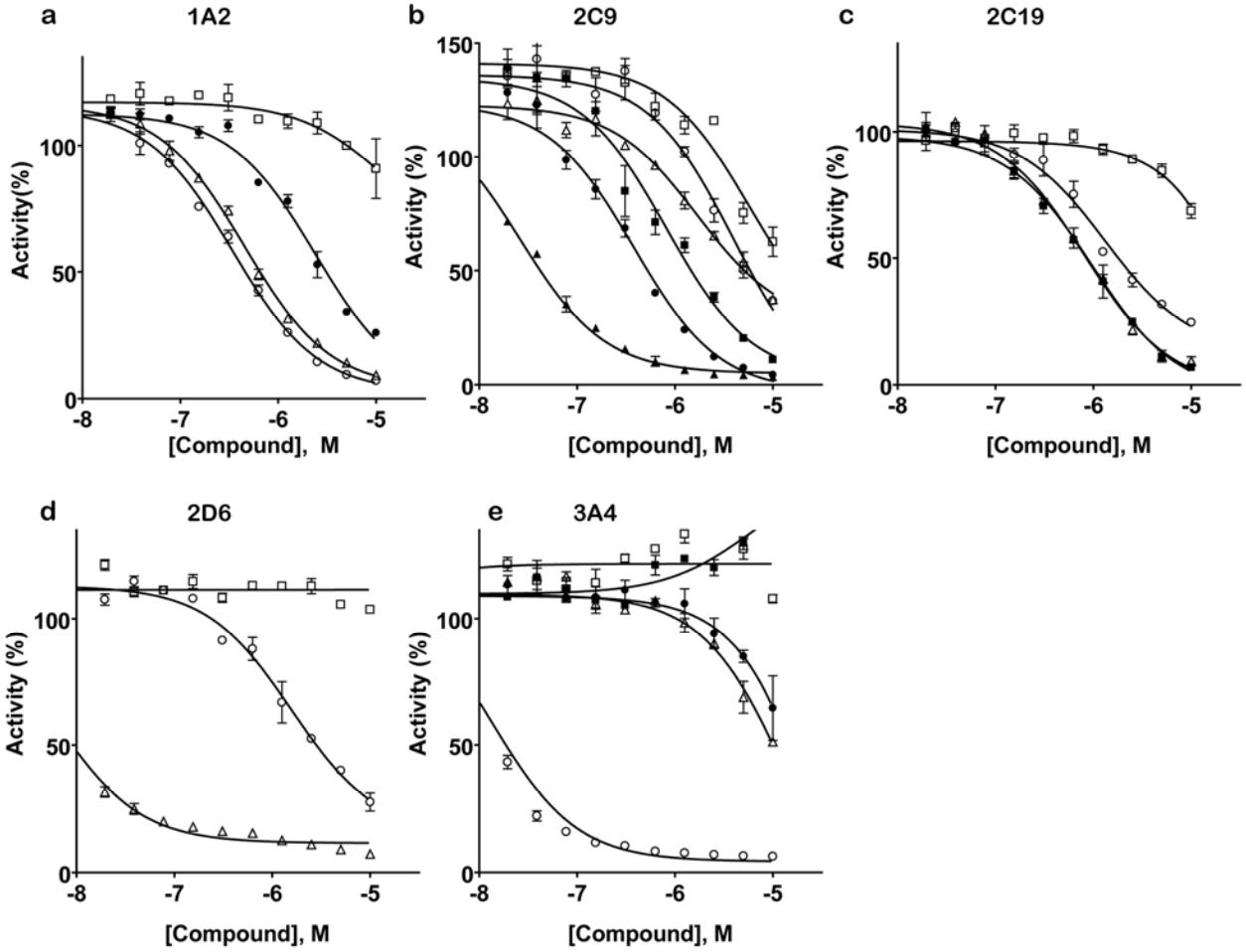
Table S1 – **Final automated 1536-well assay protocol**

Step	Parameter	Value	Description
1	Reagent	2 µL	Isozyme and Pro-Luciferin
2	Library compounds	23 nl	Compound dilution series
3	Controls	23 nl	Isozyme specific inhibitor
4	Reagent	2 µL	NADPH regeneration solution
5	Incubation time	60 min	Isozyme specific temperature
6	Reagent	4 µL	Detection Reagent
7	Incubation time	20 min	Isozyme specific temperature
8	Detection	1 sec	ViewLux luminescent read

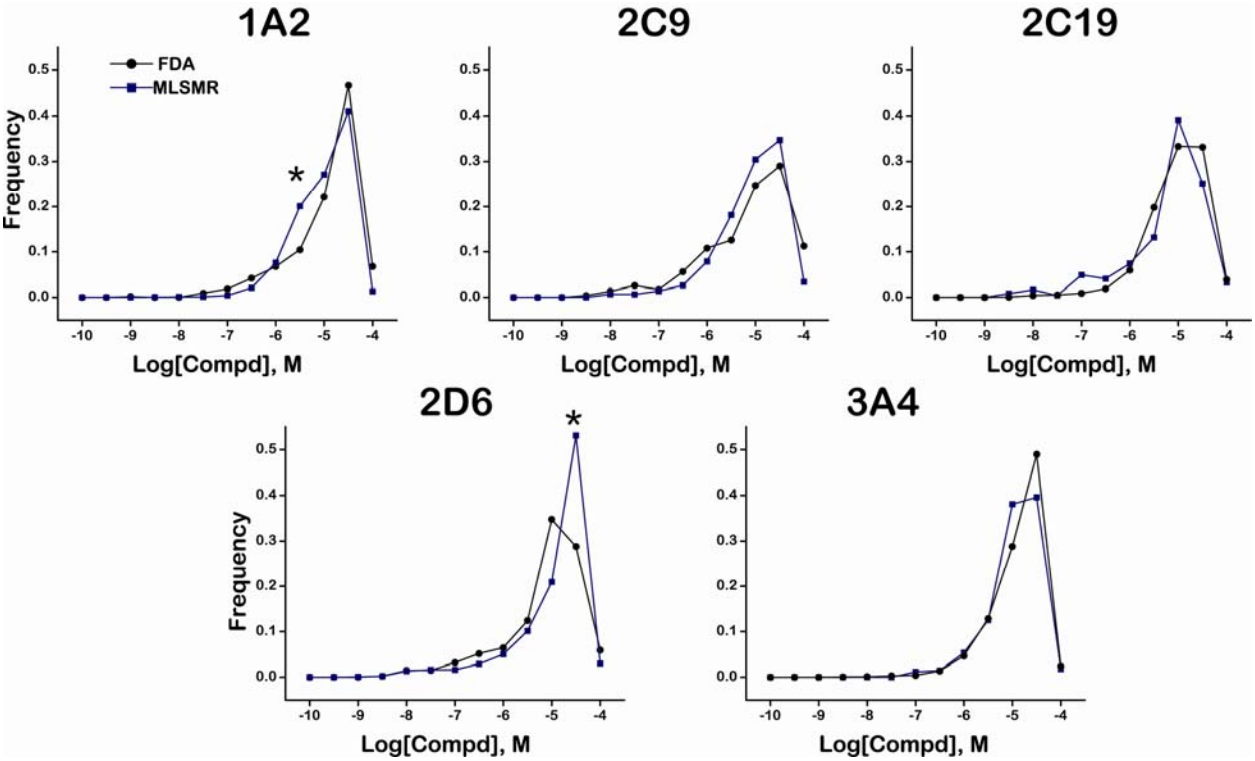
Step	Notes
1	Medium-binding white solid Kalypsys plates. Mixture kept on ice. **1A2**: 10 nM 1A2, 100 µM Luc-ME, 100 mM KPO_4; **2C9**: 10 nM 2C9, 100 µM Luc-H, 25 mM KPO_4; **2C19**: 5 nM 2C19, 10 µM Luc-H EGE, 50 mM KPO_4; **2D6**: 5 nM 2D6, 30 µM Luc-ME EGE, 100 mM KPO_4; **3A4**: 10 nM, 25 µM Luc-PPXE. All amounts are reflected in their final concentration, working concentration is 2x.
2	Pin-tool transfer compound library for a (final) range of 57 µM to 3.6 nM (7 point titration) or 0.7 nM (15 point titration)
3	Pin-tool transfer, Column 1 and 2: two-fold sixteen-point titration starting at 57 µM to 0.002 nM (final) for 1A2, 2C9, 2C19, and 3A4; 1.43 µM to 0.044 nM (final) for 2D6, Column 3: DMSO, Column 4: 57 µM (final) for 1A2, 2C9, 2C19, and 3A4; 1.43 µM for 2D6. Controls used: Furafylline (**1A2**), Sulfaphenazole (**2C9**), Ketoconazole (**2C19** and **3A4**), and Quinidine (**2D6**). Pin-tool transfer tip wash sequence: DMSO, iPA, MeOH, 3 s vacuum dry.
4	NADPH regeneration solution: 1.3 mM NADP+, 3.3 mM glucose-6-phosphate, 3.3 mM $MgCl_2$, 0.4 U/mL glucose-6-phosphate dehydrogenase for all isozymes, 200 mM KPO_4 for **3A4**. Mixture prepared at RT but kept on ice during automated run.
5	RT incubation in auxilliary plate hotel for **1A2**, **2D6**, and **3A4**; 37˚C for **2C9** and **2C19**.
6	P450 Glo-Buffer for **1A2**, **2C9**, **3A4**; Luciferin Detection Buffer for **2C19** and **2D6**. Mixture kept on ice and shielded from light.
7	RT incubation for **1A2**, **2D6**, and **3A4**; 37˚C for **2C9** and **2C19**. 60 s exposure, gain high, and 2x binning luminescent read on Perkin Elmer ViewLux.



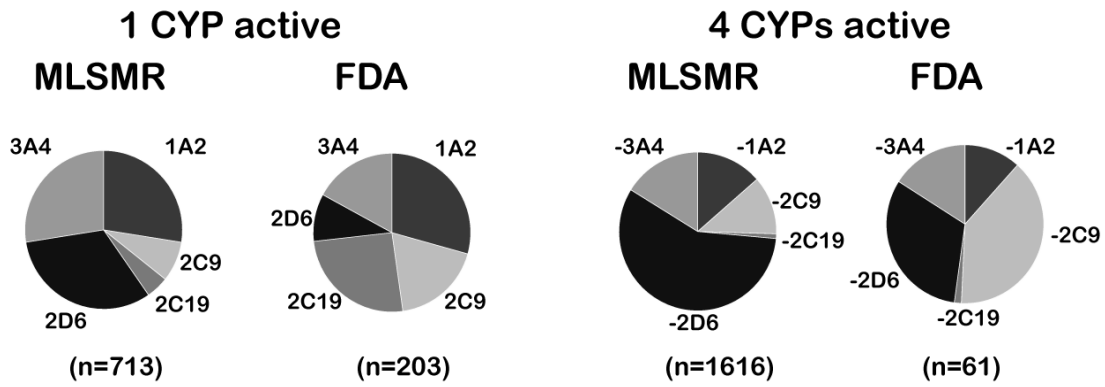
Supplementary Figure 1. Example confirmation data. Confirmation data for 91 compounds against the five CYPs **(a-e)** that were tested as 24 point two-fold dilutions titrations. Category 1 CRCs (○) and category 2 CRCs (●). Linear fits shows are of the Category 1 CRCs, the dotted line represents the 95% confidence interval. Although the CYP 2C19 **(c)** data showed the lowest Z'-factor, this dataset confirmed well upon retesting (87%). Compounds were randomly selected for confirmation from 17K data set. These compounds had also been tested by the Burnham Center for Chemical Genomics (BCCG; AIDs: 777, 778, 1024 and 1025) for CYP activity. All of the compounds published to be active by BCCG were also found as active in the present CYP assay dataset. Several additional compounds were also found to be active in the present CYP isozyme assays that were published in PubChem AIDs 1024 and 1025 to be inactive, which is likely due to the greater concentration range that was tested in the qHTS described here.



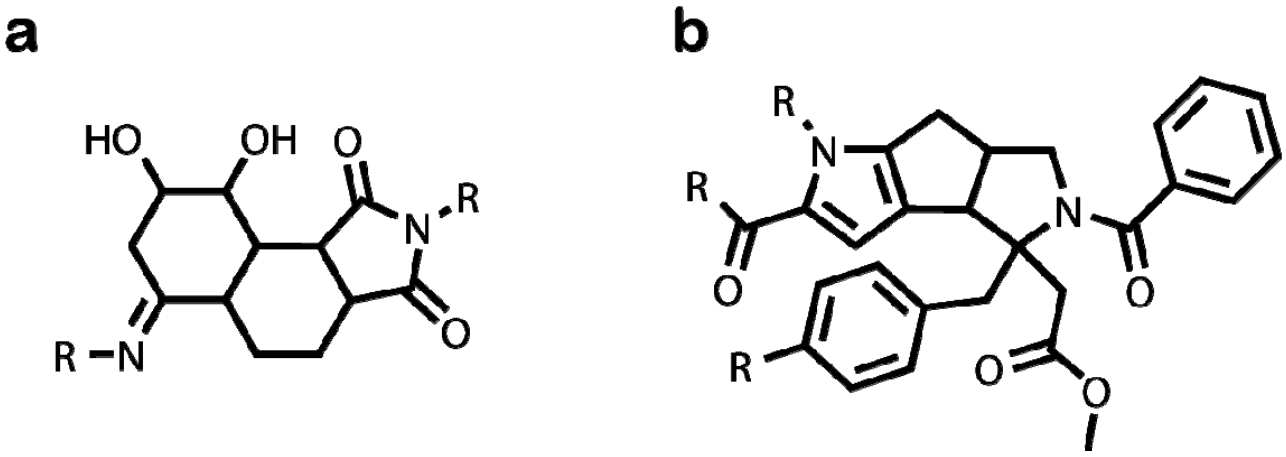
Supplementary Figure 2. Response of inhibitors and substrates in the CYP isozyme assays. Inhibitors and substrates listed as acceptable by the FDA were tested in the automated assay. The compounds are followed with the reported K_is and K_Ms. **(a)** Theophylline (□ K_M: 280-1230μM), Phenacetin (● K_M: 1.7-152μM).[40, 41] Tacrine (△ K_M: 2.8, 16μM), Furafylline (○ K_i: 0.6-0.73μM control). **(b)** Tolbutamide (□ K_M: 67-838μM), Phenytoin (○ K_M: 11.5-117μM), Flurbiprofen (△K_M: 6-42μM), Warfarin (■ K_M: 1.5-4.5μM), Diclofenac (● K_M: 3.4-52μM), Sulfaphenazole (▲ K_i: 0.3μM control). **(c)** Mephenytoin (□ K_M: 13-35μM), Fluoxetine (○ K_M: 3.7-104μM), Omeprazole (■ K_M: 17-26μM), Ketoconazole (△, control). **(d)** Debrisoquine (□ K_M: 5.6μM), Dextromethorphan (○ K_M: 0.44-8.5μM), Quinidine (△ 0.027-0.4μM control). **(e)** Testosterone (□ K_M: 52-94μM), Terfenadine (■ K_M: 15μM), Erythromycin (● K_M: 33-88μM), Nifedipine (△ K_M: 5.1-47μM), Ketoconazole (○ K_i: 0.0037-0.18μM control).



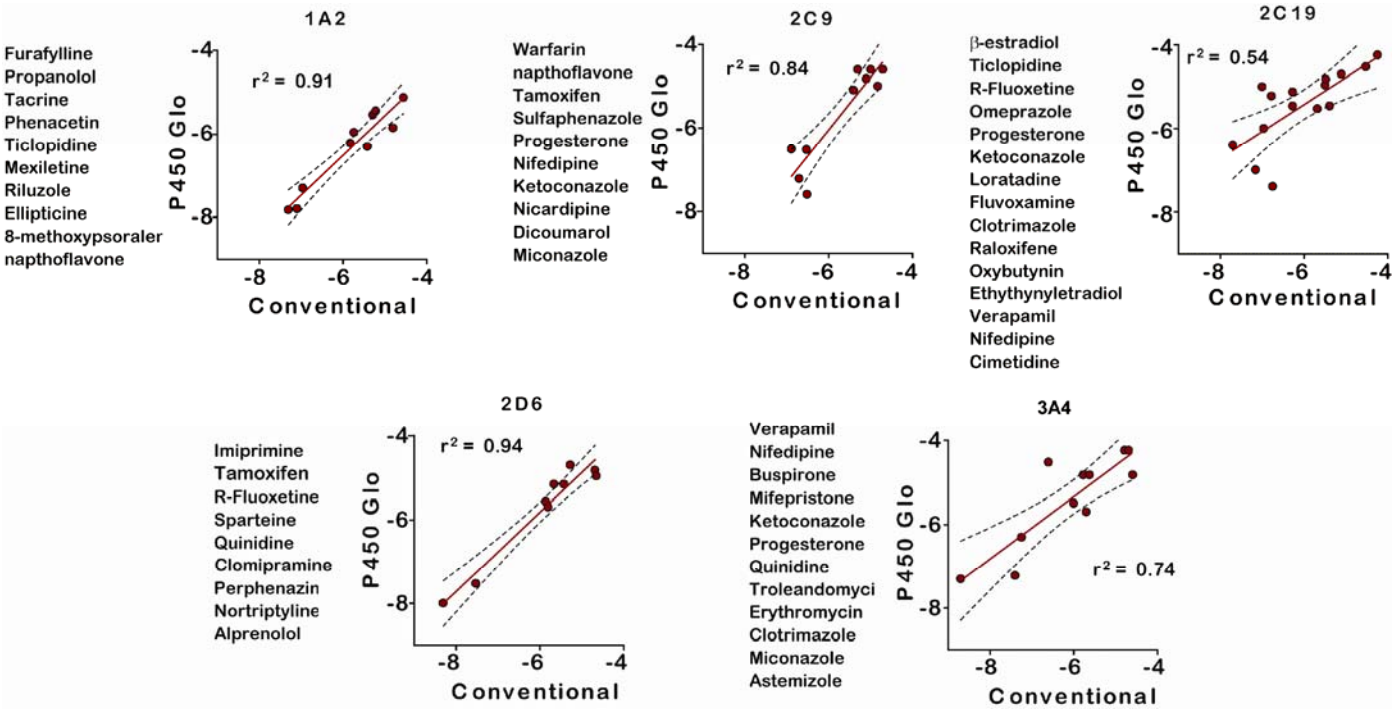
Supplementary Figure 3. Potency distribution comparison between the drug and MLSMR library sets. The * refers to p <0.01. The MLSMR is shown with a blue line and solid squares and the FDA set is shown with a black line and solid circles.



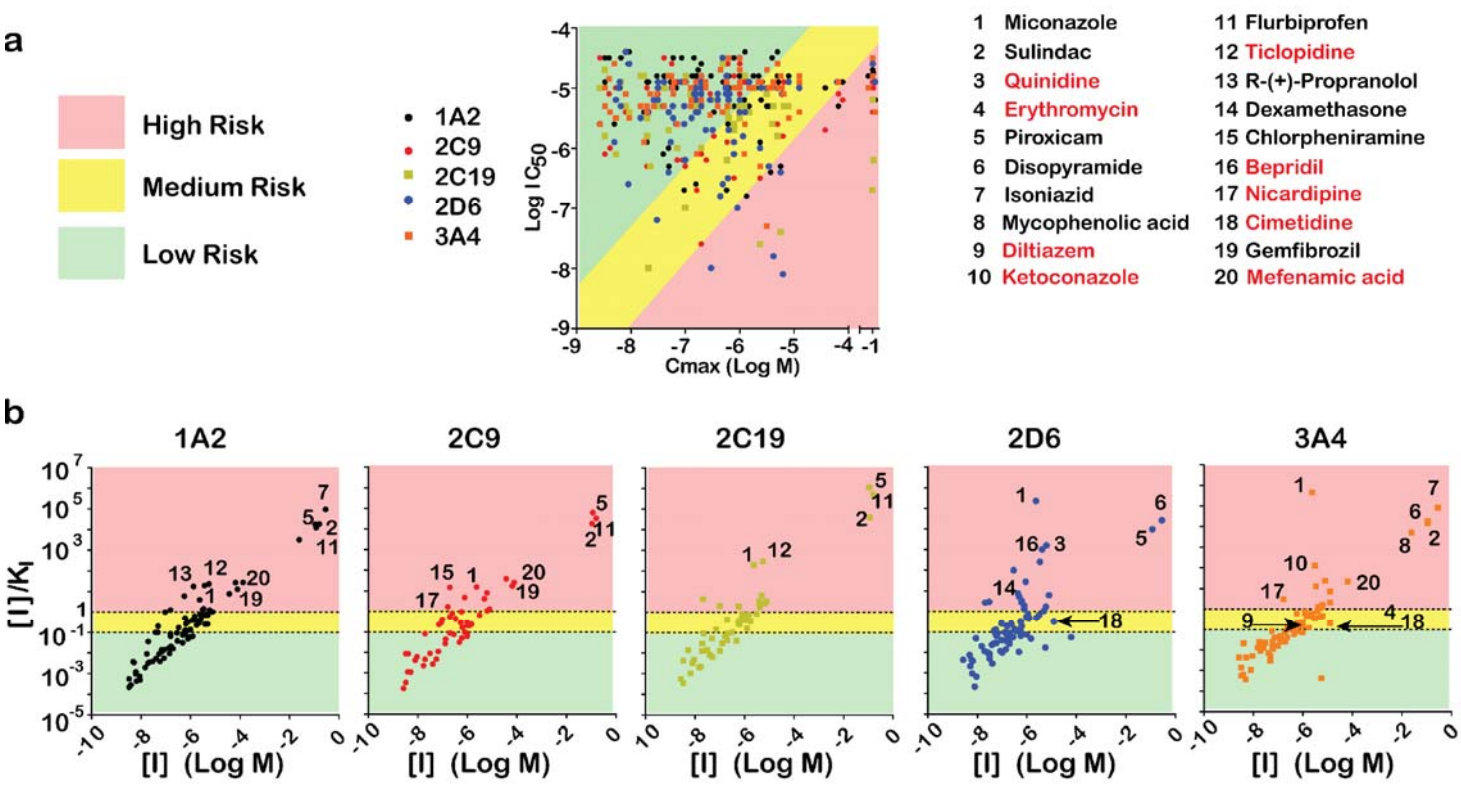
Supplementary Figure 4. Distribution of CYP isozyme activity between MLSMR and FDA sets. Pie charts represent the case when only one CYP is active for the MLSMR or FDA sets or the case when four of the five CYP isozymes were active. The active CYP is designated in the left two pie charts while the CYP that was not active is designated as for example "**-3A4**" in the right two pie charts**.** Examining the case where only one CYP isozyme is active**,** the sequence homology and the high degree of activity for isozymes CYP 2C9 and CYP 2C29 against the un-optimized MLSMR collection makes it improbable to observe these isozymes acting alone. In contrast, the FDA set showed a more uniform activity across any of the five isozymes when only one CYP isozyme was found as active with the exception of CYP 2D6. When examining the case when four isozymes were active we noted again that CYP 2C9 was often the inactive isozyme for the FDA set, supporting the selectivity of this enzyme against the FDA compounds. However, CYP 2C19 did not appear prominent when examined in context of the other isozymes, suggesting a redundant activity. Additionally, CYP 2D6, which showed the lowest activity within the biodiverse set, was often found to be the inactive isozyme in either the biodiverse or FDA set when four isozymes were active. This suggests a different selectivity for this isozyme, albeit one that does not distinguish well between the biodiverse set of compounds and the drugs.



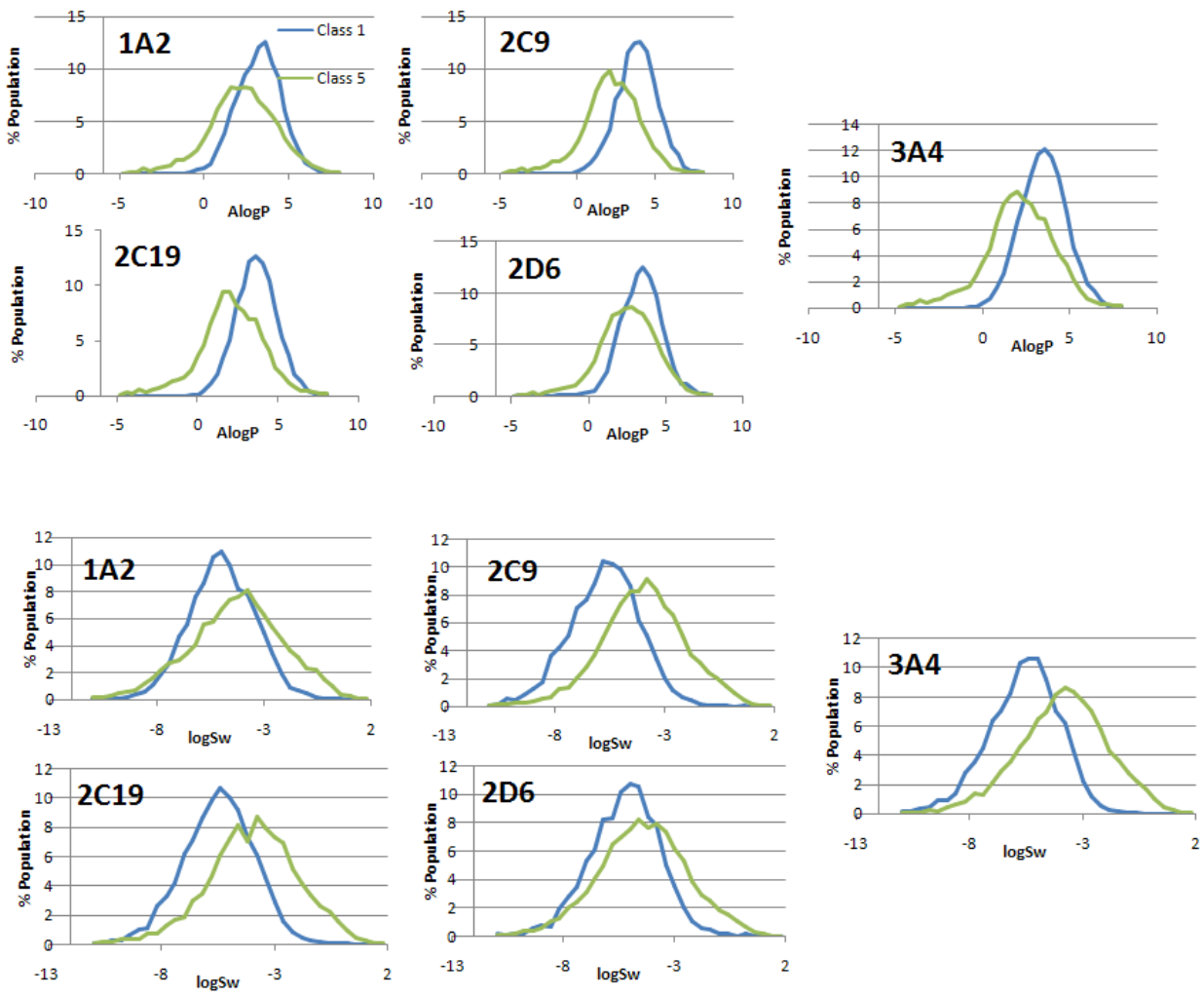
Supplementary Figure 5. Representative scaffolds for two sub-libraries.



Supplementary Figure 6. Comparison potencies for the CYP bioluminescent assays to conventional methods. The bioluminescent assay showed weaker potencies for CYP 2C19 which could be related to probe-specific behavior and suggests that the luciferin-based substrate is similar to the CYP 2C19 substrate probe S-fluoxetine. However, all the CYP 2C19 actives identified in the conventional methods were found with the bioluminescent assay.



Supplementary Figure 7. **Comparison of C_{max} plasma levels to IC_{50} obtained for the five CYP isozymes.** (**a**) Shown is the potency for each of the five CYPs plotted against the C_{max} concentration. Three regions are highlighted: Low risk, green region highlights drugs where the IC_{50} value is 10-fold above the observed plasma levels; medium risk, yellow region is where the IC_{50} value is within the observed plasma concentration; high risk red region represents drugs where the IC_{50} value is 10-fold lower than the observed plasma concentration. (**b**) The data is re-plotted for each CYP isozyme to highlight how the inhibitor affinity (K_I) compares to C_{max} ([I] is taken as the C_{max} value). For this purpose we assume a competitive inhibition model to calculate the K_I from the IC_{50}. The dotted lines highlight thresholds of low, medium and high risk regions based on FDA recommendations for the [I]/ K_I ratio. Example drugs are numbered and these are colored red if drug-drug interactions have been noted.



Supplementary Figure 8. Analysis of bulk compound properties. The population distribution between compounds showing active category 1 CRCs (blue line) or inactive (category 5, green line) were analyzed. No difference was observed between the two populations for basic compound descriptors such as MW, RB, HBA, HBD (not shown). Slight differences were observed when ALogP and LogSw were plotted and are shown here.