

## Supplemental Tables:

**Supplemental Table 1.** Confusion matrix for the training and small test set of the PXR dataset (20). The values in the parenthesis are for the small test set (20). The numbers in italics (docking and logistic regression) includes failed molecules.

### (a) Recursive Partitioning

	Activator	Non-activator
Activator	94 (14)	4 (1)
Non-activator	18	61

### (b) Random Forest

	Activator	Non-activator
Activator	81 (12)	17 (3)
Non-activator	30	49

### (c) Support Vector Machine

	Activator	Non-activator
Activator	97 (13)	1 (2)
Non-activator	9	70

### (d) Docking and logistic regression

	Activator	Non-activator
Activator	45	30, 53
Non-activator	43	29, 36

**Supplemental Table 2.** PXR confusion matrix analysis of the large 145 molecule test set. The numbers in italics (docking and logistic regression) includes failed molecules.

(a) Recursive Partitioning

	Activator	Non-activator
Activator	53	29
Non-activator	24	39

(b) Random Forest

	Activator	Non-activator
Activator	53	29
Non-activator	21	42

(c) Support Vector Machine

	Activator	Non-activator
Activator	56	26
Non-activator	22	41

(d) Docking and logistic regression

	Activator	Non-activator
Activator	58	21, <i>24</i>
Non-activator	47	10, <i>16</i>