|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Supplementary table 1. Training and test set molecules, their predicted activity and predicted error for QSAR model** | | | | | | | | | | | | | | | |
| **Reverse Transcriptase inhibitors** | | | | **Entry inhibitors** | | | | **Integrase inhibitors** | | | | **Protease inhibitors** | | | |
| **Ligand code** | **Biological Activity** | **Predicted Activity** | **Prediction Error** | **Ligand code** | **Biological Activity** | **Predicted Activity** | **Prediction Error** | **Ligand code** | **Biological Activity** | **Predicted Activity** | **Prediction Error** | **Ligand code** | **Biological Activity** | **Predicted Activity** | **Prediction Error** |
| 441300 | 6 | 5.4791 | -0.5209 | \*5100 | 4 | 4.57176 | 0.57176 | 370 | 4.14 | 4.2125 | 0.0725 | 148192 | 8.3 | 8.02 | -0.28 |
| \*50599 | 5 | 5.19182 | 0.19182 | \*5361 | 5 | 5.04342 | 0.04342 | \*6038 | 8.14 | 7.89 | -0.25 | 213039 | 7.5 | 7.45 | -0.05 |
| \*60877 | 6.1 | 6.15446 | 0.05446 | \*9064 | 5.397 | 5.0857 | -0.3113 | 23553 | 5.568 | 5.463 | -0.105 | 131536 | 7.096 | 6.27 | -0.826 |
| 60825 | 4.8 | 6.1476 | 1.3476 | 2259 | 4.823 | 5.012 | 0.189 | 235551 | 4.844 | 4.686 | -0.158 | 5362440 | 9.38 | 7.28 | -2.1 |
| 18283 | 5.3 | 6.05413 | 0.75413 | 37720 | 5.096 | 5.08208 | -0.01392 | \*271031 | 4.013 | 4.08382 | 0.07082 | 64143 | 4.65 | 4.61 | -0.04 |
| 6398764 | 5.07 | 5.19656 | 0.12656 | 64971 | 5.853 | 5.79957 | -0.05343 | 271032 | 4.113 | 4.342 | 0.229 | \*392622 | 5.2 | 5.39 | 0.19 |
| 35370 | 8.522 | 6.86827 | -1.65373 | \*65015 | 7.35 | 7.21348 | -0.13652 | 271034 | 4.627 | 4.757 | 0.13 | 441243 | 7.522 | 7.37 | -0.152 |
| \*64139 | 7.6 | 6.58187 | -1.01813 | \*65064 | 4 | 4.57224 | 0.57224 | 276805 | 4.214 | 4.8956 | 0.6816 | 54682461 | 6.38 | 6.31 | -0.07 |
| 193962 | 8.3 | 5.95781 | -2.34219 | \*460276 | 4.142 | 5.002309 | 0.860309 | \*281348 | 4.187 | 4.103 | -0.084 | 92727 | 6.99 | 6.902 | -0.088 |
| 4463 | 7 | 5.79236 | -1.20764 | 460277 | 4 | 3.88608 | -0.11392 | \*296057 | 5.055 | 4.913 | -0.142 | 65016 | 7.83 | 6.35 | -1.48 |
| 6451164 | 8.6 | 8.06301 | -0.53699 | 490397 | 5.528 | 5.57699 | 0.04899 | \*332431 | 4.096 | 4.081 | -0.015 | \*513956 | 7.3 | 6.82 | -0.48 |
| 5625 | 6.5 | 6.36486 | -0.13514 | 507806 | 8.3 | 8.20807 | -0.09193 | 364448 | 4.35 | 4.30773 | -0.04227 | 5743186 | 7.82 | 6.88 | -0.94 |
| \*90469937 | 8.522 | 7.97704 | -0.54496 | 637775 | 4.154 | 4.0934 | -0.0606 | 370661 | 4.863 | 4.88596 | 0.02296 | 461204 | 8.79 | 8.49 | -0.3 |
| 44593364 | 4.975 | 4.82337 | -0.15163 | 1570601 | 5.52 | 5.49472 | -0.02528 | 422571 | 3.95 | 4.88596 | 0.93596 | 154044 | 6.3 | 7.16 | 0.86 |
| 5464368 | 5.09 | 6.03434 | 0.94434 | 3001322 | 8.69 | 8.65309 | -0.03691 | 475513 | 6.769 | 6.69801 | -0.07099 | \*21696352 | 6.72 | 7.02 | 0.3 |
| \*16129720 | 5.301 | 5.23083 | -0.07017 | \*3002977 | 8.34 | 8.71664 | 0.37664 | 475514 | 5.455 | 5.49703 | 0.04203 | \*54678486 | 4.522 | 4.37 | -0.152 |
| 463247 | 6.04 | 5.43326 | -0.60674 | 3009355 | 9.04 | 9.14171 | 0.10171 | 3000873 | 4.721 | 4.67044 | -0.05056 | \*44456632 | 7.301 | 7.01 | -0.291 |
| 40304 | 4.657 | 5.02866 | 0.37166 | 5279270 | 5.376 | 5.49472 | 0.11872 | 3000874 | 4.853 | 4.92826 | 0.07526 | 16072580 | 6.782 | 6.53 | -0.252 |
| \*25659 | 4.721 | 4.84247 | 0.12147 | 5280343 | 4.69 | 5.08676 | 0.39676 | \*5276374 | 4.69 | 4.46038 | -0.22962 | 446837 | 7.283 | 7.126 | -0.157 |
| 72323 | 3.946 | 4.1837 | 0.2377 | 5280863 | 5.397 | 5.15855 | -0.23845 | 5277135 | 9.15 | 9.13375 | -0.01625 | \*25058165 | 8.301 | 8.503 | 0.202 |
| 12456 | 3.785 | 3.81169 | 0.02669 | 5281607 | 4.69 | 4.80158 | 0.11158 | 5280343 | 4.958 | 4.94706 | -0.01094 | 5462355 | 5.853 | 5.96 | 0.107 |
| \*72321 | 3.55 | 4.26048 | 0.71048 | 5281672 | 5.69 | 5.40556 | -0.28444 | 5281787 | 4.723 | 4.68981 | -0.03319 | 3013165 | 5.744 | 6.31 | 0.566 |
| \*5281541 | 3.397 | 3.92325 | 0.52625 | \*5359407 | 4.9 | 5.19604 | 0.29604 | 5281801 | 5.091 | 5.00822 | -0.08278 | 469665 | 7.455 | 7.856 | 0.401 |
| 72412 | 7.42 | 6.60085 | -0.81915 | 5495818 | 5.36 | 6.05717 | 0.69717 | \*5281813 | 5.397 | 5.2743 | -0.1227 | \*73096 | 7.958 | 8.11 | 0.152 |
| \*455249 | 6.88 | 5.80959 | -1.07041 | 6364572 | 5.69 | 5.7825 | 0.0925 | 5287411 | 6.187 | 6.2811 | 0.0941 | 5480679 | 7.95 | 8.57 | 0.62 |
| 107905 | 10.07 | 9.32877 | -0.74123 | 6510265 | 6.096 | 6.07967 | -0.01633 | 9574291 | 4.018 | 4.05029 | 0.03229 | 158550 | 5.229 | 6.47 | 1.241 |
| 5280863 | 3.496 | 4.88287 | 1.38687 | \*9566057 | 5.42 | 6.034416 | 0.614416 | 11542423 | 8 | 7.98011 | -0.01989 | 131535 | 7.096 | 6.803 | -0.293 |
| 5280343 | 3.204 | 4.6853 | 1.4813 | 10010220 | 4.508 | 4.5475 | 0.0395 | 20744900 | 7.69 | 7.64579 | -0.04421 | \*5281813 | 7.079 | 7.135 | 0.056 |
| 30323 | 3.625 | 3.11515 | -0.50985 | 10010385 | 4.53 | 4.71075 | 0.18075 | 54606068 | 6.537 | 6.50271 | -0.03429 | 121285264 | 9.096 | 8.017 | -1.079 |
| 460509 | 4.408 | 4.18356 | -0.22444 | 11285792 | 8.22 | 8.15993 | -0.06007 | \*54671008 | 7.39 | 7.21815 | -0.17185 | 60927 | 5.187 | 5.371 | 0.184 |
| 341 | 3.698 | 5.88766 | 2.18966 | 11319217 | 7.88 | 7.95345 | 0.07345 | 54682040 | 7.82 | 7.81771 | -0.00229 | \*65012 | 8.638 | 8.535 | -0.103 |
| 14483070 | 7.301 | 9.50456 | 2.20356 | 13892291 | 5.408 | 5.33174 | -0.07626 | \*54689520 | 6 | 6.23129 | 0.23129 | \*3033932 | 8.552 | 8.786 | 0.234 |
| 610378 | 5.528 | 5.17555 | -0.35245 | 16130199 | 8.25 | 8.23929 | -0.01071 | 54698642 | 8.58 | 8.56305 | -0.01695 | 455550 | 8.823 | 8.792 | -0.031 |
| 3000870 | 7.823 | 7.53 | -0.293 | 16130283 | 4.69 | 4.68593 | -0.00407 | 54706524 | 7.29 | 7.32506 | 0.03506 | 5464035 | 9.455 | 10.02 | 0.565 |
| 16045418 | 9.4 | 8.597 | -0.803 | 16130418 | 5.36 | 5.34725 | -0.01275 | 54713659 | 8.5 | 8.52192 | 0.02192 | 11765769 | 8.3 | 9 | 0.7 |
| 468856 | 5.79 | 5.8 | 0.01 | \*16134407 | 5.698 | 5.84526 | 0.14726 | 54726191 | 8.56 | 8.58782 | 0.02782 | 66868 | 6.01 | 6.45 | 0.44 |
| 40304 | 5 | 4.96 | -0.04 | 16729392 | 6.97 | 6.22261 | -0.74739 | \*66561902 | 7.82 | 7.45804 | -0.36196 | 72404 | 7.447 | 8.025 | 0.578 |
| 5281541 | 4.346 | 5.342 | 0.996 | 53629732 | 4.69 | 4.71244 | 0.02244 | 71308208 | 6.19 | 6.1935 | 0.0035 | 72290 | 7.341 | 7.451 | 0.11 |
| 3008867 | 6.29 | 6.15 | -0.14 | 124081059 | 4 | 4.06785 | 0.06785 | 71589630 | 5.6 | 5.59758 | -0.00242 | 65126 | 7.356 | 7.63 | 0.274 |
| \*477596 | 5.958 | 5.63 | -0.328 | 124203200 | 5.2 | 5.22559 | 0.02559 | 90311989 | 8.12 | 8.16843 | 0.04843 | 461207 | 6.903 | 7.842 | 0.939 |

\* Test set and remaining is training set