Supplementary Information for Chemformer: A Pre-Trained Transformer for Computational Chemistry

Ross Irwin¹, Spyridon Dimitriadis¹,², Jiazhen He¹, and Esben Jannik Bjerrum¹,*

1) Molecular AI, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden
2) Department of Computer and Information Science, Linköping University, Linköping, Sweden
*) esben.bjerrum@astrazeneca.com
SI table 1: Identified Hyperparameters for biological activity models

| Hyperparameter          | Randomly initialised | Pre-Trained Chemformer |
|-------------------------|----------------------|------------------------|
| Number of Encoder layers| 4                    | 6                      |
| Model dimension (d_model)| 512                  | 512                    |
| Number of attention heads| 8                    | 8                      |
| Encoder FNN size        | 512                  | 2048                   |
| Dropout for encoder     | 0.14                 | 0.15                   |
| Regression FNN size     | [1024, 512, 1]       | [1024, 512, 1]         |
| Dropout for regression FNN| 0.62                 | 0.6                    |
| Max Learning rate       | 0.0002               | 0.0003                 |
| Weight decay            | 0.0000007            | 0                      |

SI table 2: Identified Hyperparameters for Property Prediction Models

| Hyperparameter          | Randomly initialised | Pre-Trained Chemformer |
|-------------------------|----------------------|------------------------|
| Number of Encoder layers| 2                    | 6                      |
| Model dimension (d_model)| 256                  | 512                    |
| Number of attention heads| 8                    | 8                      |
| Encoder FNN size        | 1024                 | 2048                   |
| Dropout for encoder     | 0.15                 | 0.17                   |
| Regression FNN size     | [1024, 1]            | [1024, 1]              |
| Dropout for regression FNN| 0.25                 | 0.25                   |
| Max Learning rate       | 0.0003               | 0.0005                 |
| Weight decay            | 0                    | 0                      |
| Gene Symbol | RMSE_train | R2_train | RMSE_test | R2_test |
|-------------|------------|----------|-----------|---------|
| ABL1        | 0.462419   | 0.759587 | 0.697949  | 0.500608|
| ACH1        | 0.631094   | 0.705351 | 0.875118  | 0.465458|
| ADAM17      | 0.568146   | 0.723813 | 0.754872  | 0.559902|
| ADORA2A     | 0.63258    | 0.680259 | 0.879744  | 0.439787|
| ADORA2B     | 0.592952   | 0.657747 | 0.707185  | 0.511975|
| ADORA3      | 0.689681   | 0.65958  | 0.908645  | 0.435202|
| ADRA1A      | 0.640709   | 0.745833 | 0.87532   | 0.490602|
| ADRA1D      | 0.569135   | 0.776011 | 0.885185  | 0.531825|
| ADRB1       | 0.568706   | 0.741632 | 0.792097  | 0.535432|
| ADRB2       | 0.623623   | 0.826672 | 0.991199  | 0.534796|
| ADRB3       | 0.571198   | 0.750505 | 0.802513  | 0.519703|
| AKT1        | 0.528964   | 0.82676  | 0.75026   | 0.679908|
| AKT2        | 0.405318   | 0.860979 | 0.513022  | 0.755294|
| ALK         | 0.452165   | 0.860814 | 0.782426  | 0.553251|
| ALOX5       | 0.526221   | 0.556771 | 0.672976  | 0.317122|
| AR          | 0.634198   | 0.714648 | 0.953382  | 0.428502|
| AURKA       | 0.505814   | 0.762532 | 0.737328  | 0.541808|
| AURKB       | 0.501806   | 0.79251  | 0.756298  | 0.578423|
| BACE1       | 0.599853   | 0.672238 | 0.748653  | 0.463619|
| CA1         | 0.596647   | 0.678075 | 0.742796  | 0.472154|
| CA12        | 0.540507   | 0.740004 | 0.689447  | 0.6312  |
| CA2         | 0.585445   | 0.654    | 0.684393  | 0.54418 |
| CA9         | 0.569501   | 0.702427 | 0.677287  | 0.58182 |
| CASP1       | 0.476532   | 0.867736 | 0.686196  | 0.705846|
| CCKBR       | 0.726013   | 0.664643 | 1.04335   | 0.39449 |
| CCR2        | 0.64064    | 0.62906  | 0.852511  | 0.396741|
| CCR5        | 0.620441   | 0.730992 | 0.84619   | 0.484079|
| CDK1        | 0.48318    | 0.760328 | 0.712554  | 0.535366|
| CDK2        | 0.516791   | 0.791409 | 0.797456  | 0.536452|
| CHEK1       | 0.561343   | 0.826775 | 0.897522  | 0.568051|
| CHRM1       | 0.649107   | 0.74454  | 1.019677  | 0.435387|
| CHRM2       | 0.614097   | 0.779785 | 0.87732   | 0.555708|
| CHRM3       | 0.616755   | 0.825438 | 0.936636  | 0.640562|
| CHRNA7      | 0.682457   | 0.659339 | 0.939164  | 0.323489|
| CLK4        | 0.558478   | 0.658439 | 0.847059  | 0.343265|
| CNR1        | 0.664118   | 0.687004 | 0.87218   | 0.490499|
| CNR2        | 0.634618   | 0.705911 | 0.88034   | 0.461537|
| CRHR1       | 0.644334   | 0.574202 | 0.790114  | 0.343007|
| CSFR1       | 0.487445   | 0.848536 | 0.743496  | 0.652707|
| CTSDK       | 0.717651   | 0.707896 | 0.995487  | 0.417651|
| CTSS        | 0.560175   | 0.745725 | 0.820412  | 0.479351|
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| CYP19A1| 0.646602| 0.673974| 1.00146 | 0.29631 |
| DHFR   | 0.833097| 0.707825| 1.14686 | 0.43637 |
| DPP4   | 0.610023| 0.636442| 0.788579| 0.460189|
| DRD1   | 0.669167| 0.650512| 0.962922| 0.384425|
| DRD3   | 0.658044| 0.722821| 0.939136| 0.447912|
| DRD4   | 0.649427| 0.613144| 0.88415 | 0.301673|
| DYRK1A | 0.520324| 0.604236| 0.654494| 0.375898|
| EDNRA  | 0.699812| 0.731035| 0.901545| 0.536975|
| EGFR   | 0.643551| 0.718310| 0.909381| 0.465569|
| EPHX2  | 0.676169| 0.59836 | 0.898318| 0.281056|
| ERBB2  | 0.461802| 0.776674| 0.648461| 0.58203 |
| ESR1   | 0.640037| 0.733328| 0.937766| 0.476277|
| ESR2   | 0.64227 | 0.675607| 0.936112| 0.343888|
| F10    | 0.655742| 0.776049| 0.955131| 0.561916|
| F2     | 0.682981| 0.745513| 1.000961| 0.499757|
| FAAH   | 0.665409| 0.704811| 0.900961| 0.499757|
| FGFR1  | 0.409379| 0.807712| 0.588183| 0.619093|
| FLT1   | 0.505066| 0.75158 | 0.671265| 0.582197|
| FLT3   | 0.514917| 0.802938| 0.782282| 0.557507|
| GHSR   | 0.64556 | 0.624921| 0.897273| 0.258657|
| GNRHR  | 0.724431| 0.635148| 0.868312| 0.471693|
| GRM5   | 0.648181| 0.57567 | 0.862893| 0.332537|
| GSK3A  | 0.435357| 0.770372| 0.677902| 0.459463|
| GSK3B  | 0.558298| 0.722173| 0.765797| 0.482102|
| HDAC1  | 0.542348| 0.701989| 0.8002  | 0.404979|
| HPGD   | 0.388397| 0.350666| 0.465574| 0.265302|
| HRH3   | 0.642065| 0.644979| 0.821974| 0.452561|
| HSD11B1| 0.635596| 0.598246| 0.833858| 0.377406|
| HSP90AA1| 0.417647| 0.814345| 0.66158 | 0.54519 |
| HTR2A  | 0.656454| 0.699693| 0.928723| 0.427859|
| HTR2C  | 0.632339| 0.621365| 0.882977| 0.288048|
| HTR6   | 0.658896| 0.662438| 0.898682| 0.394009|
| HTR7   | 0.626396| 0.623255| 0.878066| 0.324537|
| IGF1R  | 0.444245| 0.792011| 0.52696 | 0.713071|
| INSR   | 0.391834| 0.753515| 0.479435| 0.596118|
| ITK    | 0.459493| 0.815336| 0.698489| 0.515192|
| JAK2   | 0.497192| 0.828371| 0.774993| 0.578 |
| JAK3   | 0.424532| 0.780655| 0.575762| 0.562296|
| KCNH2  | 0.500187| 0.605239| 0.694129| 0.284215|
| KDR    | 0.566472| 0.749058| 0.845926| 0.495487|
| KIT    | 0.506487| 0.766946| 0.726072| 0.502735|
| LCK    | 0.496114| 0.816935| 0.754742| 0.579422|
| MAOB   | 0.70377 | 0.593893| 0.937241| 0.293393|
| MAPK14 | 0.578347| 0.730236| 0.830582| 0.448016|
| MAPK8  | 0.426594| 0.767619| 0.667562| 0.536197|
| MAPK9  | 0.406851| 0.77268 | 0.578377| 0.542752|
| Gene    | Expression 1 | Expression 2 | Expression 3 | Expression 4 |
|---------|-------------|-------------|-------------|-------------|
| MAPKAPK2| 0.41606     | 0.839766    | 0.564356    | 0.707083    |
| MC4R    | 0.675689    | 0.622051    | 0.797481    | 0.480612    |
| MCHR1   | 0.60605     | 0.636705    | 0.875276    | 0.329054    |
| MET     | 0.48029     | 0.815941    | 0.704251    | 0.640846    |
| MMP1    | 0.595042    | 0.682163    | 0.775696    | 0.439851    |
| MMP13   | 0.619178    | 0.731744    | 0.915197    | 0.428035    |
| MMP2    | 0.658661    | 0.759502    | 0.976917    | 0.480498    |
| MMP3    | 0.611682    | 0.66531     | 0.871344    | 0.380656    |
| MMP9    | 0.656034    | 0.697322    | 0.95119     | 0.409055    |
| MTOR    | 0.6247      | 0.784053    | 0.63066     | 0.66142     |
| NPY5R   | 0.657255    | 0.622725    | 0.869299    | 0.324074    |
| NR3C1   | 0.631136    | 0.718529    | 0.886694    | 0.463159    |
| NTRK1   | 0.397897    | 0.817824    | 0.63066     | 0.473491    |
| OPRD1   | 0.796158    | 0.66134     | 1.098622    | 0.396241    |
| OPRK1   | 0.683332    | 0.770837    | 0.960594    | 0.563126    |
| OPRL1   | 0.629091    | 0.684508    | 0.890089    | 0.426156    |
| OPNM1   | 0.758512    | 0.714537    | 1.057048    | 0.473367    |
| P2RX7   | 0.574104    | 0.523276    | 0.784893    | 0.223285    |
| PARP1   | 0.558386    | 0.684444    | 0.827246    | 0.391962    |
| PDE5A   | 0.630056    | 0.814421    | 0.905679    | 0.636527    |
| PDGFRB  | 0.496985    | 0.796679    | 0.825897    | 0.499163    |
| PGR     | 0.591037    | 0.768732    | 0.863572    | 0.542742    |
| PIK3CA  | 0.60058     | 0.749063    | 0.8017      | 0.533754    |
| PIM1    | 0.466648    | 0.85467     | 0.677689    | 0.716151    |
| PIM2    | 0.362492    | 0.858067    | 0.637732    | 0.635243    |
| PLK1    | 0.367063    | 0.760219    | 0.493136    | 0.580475    |
| PPARA   | 0.665272    | 0.584221    | 0.910217    | 0.28742     |
| PPARD   | 0.593576    | 0.73638     | 0.885797    | 0.444862    |
| PPARG   | 0.598302    | 0.696012    | 0.857019    | 0.408688    |
| PRKACA  | 0.418943    | 0.717004    | 0.539586    | 0.545742    |
| PRKCD   | 0.462696    | 0.841132    | 0.649857    | 0.666639    |
| PTGDR2  | 0.650644    | 0.620702    | 0.884769    | 0.337231    |
| PTGS2   | 0.658691    | 0.618428    | 0.815599    | 0.462617    |
| PTPN1   | 0.405994    | 0.759641    | 0.52879     | 0.562502    |
| REN     | 0.752667    | 0.67535     | 0.977488    | 0.471566    |
| ROCK1   | 0.498509    | 0.768183    | 0.698079    | 0.537549    |
| ROCK2   | 0.555124    | 0.785698    | 0.850822    | 0.502342    |
| S1PR1   | 0.679563    | 0.754395    | 1.00735     | 0.473819    |
| SCN9A   | 0.659754    | 0.471669    | 0.786117    | 0.311671    |
| SIGMAR1 | 0.688565    | 0.595456    | 0.897492    | 0.364662    |
| SLC6A2  | 0.658597    | 0.612711    | 0.856543    | 0.380274    |
| SLC6A3  | 0.660008    | 0.650841    | 0.849639    | 0.41966     |
| SRC     | 0.547082    | 0.724974    | 0.711809    | 0.529486    |
| TACR1   | 0.66865     | 0.739477    | 0.923579    | 0.512668    |
| TRPV1   | 0.649489    | 0.680386    | 0.883044    | 0.435926    |
| VDR     | 0.514416    | 0.801084    | 0.654198    | 0.683171    |
| Gene Symbol | RMSE_train | R²_train | RMSE_test | R²_test |
|-------------|------------|----------|-----------|---------|
| ABL1        | 0.322417   | 0.882635 | 0.581919  | 0.65185 |
| ACHE        | 0.451184   | 0.849527 | 0.782727  | 0.569346|
| ADAM17      | 0.38213    | 0.874636 | 0.653015  | 0.671288|
| ADORA2A     | 0.431567   | 0.857526 | 0.725238  | 0.611128|
| ADORA2B     | 0.382286   | 0.857343 | 0.65158   | 0.596088|
| ADORA3      | 0.453086   | 0.852804 | 0.79538   | 0.562425|
| ADRA1A      | 0.433057   | 0.883524 | 0.802774  | 0.582804|
| ADRA1D      | 0.422446   | 0.877086 | 0.781807  | 0.632109|
| ADRB1       | 0.400135   | 0.871915 | 0.714793  | 0.611245|
| ADRB2       | 0.406192   | 0.92642  | 0.809468  | 0.677627|
| ADRB3       | 0.394303   | 0.880755 | 0.725398  | 0.604263|
| AKT1        | 0.36982    | 0.915348 | 0.646801  | 0.761862|
| AKT2        | 0.270657   | 0.93784  | 0.417703  | 0.836325|
| ALK         | 0.319176   | 0.9305   | 0.636381  | 0.704573|
| ALOX5       | 0.375183   | 0.774598 | 0.600432  | 0.443171|
| AR          | 0.455767   | 0.852528 | 0.890439  | 0.505809|
| AURKA       | 0.353652   | 0.883522 | 0.610824  | 0.676045|
| AURKB       | 0.332581   | 0.908761 | 0.602813  | 0.729063|
| BACE1       | 0.470828   | 0.797997 | 0.72276   | 0.504834|
| CA1         | 0.384078   | 0.866457 | 0.578599  | 0.676609|
| CA12        | 0.31767    | 0.910225 | 0.509057  | 0.79901 |
| CA2         | 0.359764   | 0.868613 | 0.562676  | 0.690308|
| CA9         | 0.371327   | 0.873198 | 0.580713  | 0.69544 |
| CASP1       | 0.345373   | 0.93037  | 0.6493    | 0.732676|
| CCKBR       | 0.541631   | 0.813235 | 0.85412   | 0.583724|
| CCR2        | 0.49191    | 0.780969 | 0.752117  | 0.523207|
| CCR5        | 0.438579   | 0.865115 | 0.822586  | 0.515796|
| CDK1        | 0.304588   | 0.904624 | 0.606796  | 0.661424|
| CDK2        | 0.340355   | 0.909477 | 0.636214  | 0.694033|
| CHEK1       | 0.351135   | 0.932181 | 0.681083  | 0.744676|
| CHRM1       | 0.429847   | 0.888088 | 0.89245   | 0.565461|
| CHRM2       | 0.432852   | 0.890361 | 0.789738  | 0.635607|
| CHRM3       | 0.4439     | 0.90954  | 0.761144  | 0.758095|
| CHRNA7      | 0.524318   | 0.798995 | 0.909163  | 0.387424|
| CLK4        | 0.355909   | 0.861313 | 0.707829  | 0.527284|
| CNR1        | 0.459961   | 0.849807 | 0.74373   | 0.630308|
| CNR2        | 0.447029   | 0.853936 | 0.755726  | 0.598938|
| CRHR1       | 0.440791   | 0.800538 | 0.691997  | 0.500233|
| CSF1R       | 0.359865   | 0.917313 | 0.641944  | 0.738541|
| CTSK        | 0.466249   | 0.875345 | 0.817106  | 0.600322|
| CTSS        | 0.395085   | 0.87333  | 0.733408  | 0.580638|
| CYP19A1     | 0.448196   | 0.843217 | 0.885008  | 0.428662|
| Gene   | MAF   | Coverage | %Mapp | %Dels |
|--------|-------|----------|-------|-------|
| DHFR   | 0.61152 | 0.842449 | 0.990868 | 0.575547 |
| DPP4   | 0.427706 | 0.82072 | 0.71247 | 0.555396 |
| DRD1   | 0.473664 | 0.824472 | 0.867415 | 0.498216 |
| DRD3   | 0.447757 | 0.871665 | 0.797763 | 0.589688 |
| DRD4   | 0.446622 | 0.81663 | 0.774734 | 0.454603 |
| Dyrk1A | 0.337753 | 0.8332 | 0.569529 | 0.530155 |
| EDNRa  | 0.504937 | 0.859732 | 0.835738 | 0.610335 |
| EGFR   | 0.431544 | 0.873238 | 0.808927 | 0.574909 |
| EPRX2  | 0.460128 | 0.813368 | 0.822018 | 0.415314 |
| ERBB2  | 0.337753 | 0.8332 | 0.569529 | 0.530155 |
| ESR1   | 0.41159 | 0.889683 | 0.756207 | 0.65625 |
| ESR2   | 0.432441 | 0.852944 | 0.81077 | 0.490422 |
| F10    | 0.463566 | 0.887938 | 0.833296 | 0.665968 |
| F2     | 0.436649 | 0.895911 | 0.812822 | 0.664436 |
| FAAH   | 0.478457 | 0.847 | 0.86441 | 0.516634 |
| FGFR1  | 0.305864 | 0.892785 | 0.533467 | 0.686664 |
| FLT1   | 0.337401 | 0.889142 | 0.648392 | 0.614213 |
| FLT3   | 0.340633 | 0.913603 | 0.73286 | 0.61218 |
| Ghsr   | 0.490624 | 0.782666 | 0.813341 | 0.372329 |
| GNRHR  | 0.619296 | 0.73318 | 0.790289 | 0.562417 |
| GRM5   | 0.436011 | 0.807791 | 0.800821 | 0.435626 |
| Gsk3A  | 0.274064 | 0.908747 | 0.509331 | 0.662052 |
| Gsk3B  | 0.338493 | 0.897682 | 0.648471 | 0.619622 |
| HAC1   | 0.366403 | 0.863654 | 0.685339 | 0.550749 |
| HPGD   | 0.299809 | 0.609256 | 0.451079 | 0.319188 |
| HRH3   | 0.465613 | 0.812882 | 0.718078 | 0.578636 |
| HSD11B1| 0.493007 | 0.757279 | 0.772447 | 0.462748 |
| HSP90AA1| 0.299064 | 0.904741 | 0.492559 | 0.736289 |
| HTR2A  | 0.436412 | 0.866952 | 0.800186 | 0.572987 |
| HTR2C  | 0.425188 | 0.828424 | 0.750789 | 0.468912 |
| HTR6   | 0.420631 | 0.861993 | 0.718981 | 0.584151 |
| HTR7   | 0.404119 | 0.842505 | 0.785253 | 0.462912 |
| IGFR1  | 0.311534 | 0.893735 | 0.453714 | 0.789418 |
| INSR   | 0.272027 | 0.880672 | 0.42104 | 0.693375 |
| ITK    | 0.306257 | 0.917361 | 0.541347 | 0.695959 |
| JAK2   | 0.333187 | 0.922705 | 0.69552 | 0.657933 |
| JAK3   | 0.317621 | 0.877011 | 0.57496 | 0.570471 |
| KCNH2  | 0.332192 | 0.82592 | 0.612833 | 0.442787 |
| KDR    | 0.398353 | 0.875869 | 0.717632 | 0.631616 |
| KIT    | 0.343636 | 0.891988 | 0.67203 | 0.567774 |
| LCK    | 0.337633 | 0.915092 | 0.633633 | 0.696541 |
| MAOB   | 0.469333 | 0.819199 | 0.882616 | 0.392144 |
| MAPK14 | 0.387764 | 0.878459 | 0.715785 | 0.581032 |
| MAPK8  | 0.286031 | 0.895271 | 0.519198 | 0.705041 |
| MAPK9  | 0.257285 | 0.908532 | 0.456173 | 0.704668 |
| MAPKAP2| 0.309104 | 0.911221 | 0.474278 | 0.792193 |
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| MC4R   | 0.522064| 0.773978| 0.736684| 0.560998|
| MCHR1  | 0.435248| 0.812375| 0.732427| 0.516777|
| MET    | 0.319847| 0.917944| 0.582309| 0.751984|
| MMP1   | 0.388372| 0.864599| 0.649067| 0.609373|
| MMP13  | 0.415963| 0.878652| 0.74618  | 0.606784|
| MMP2   | 0.406122| 0.908653| 0.788691| 0.661432|
| MMP3   | 0.407324| 0.830681| 0.743188| 0.543075|
| MMP9   | 0.396647| 0.889226| 0.776194| 0.597425|
| MTOR   | 0.319847| 0.917944| 0.582309| 0.751984|
| MMP1   | 0.388372| 0.864599| 0.649067| 0.609373|
| MMP13  | 0.415963| 0.878652| 0.74618  | 0.606784|
| MMP2   | 0.406122| 0.908653| 0.788691| 0.661432|
| MMP3   | 0.407324| 0.830681| 0.743188| 0.543075|
| MMP9   | 0.396647| 0.889226| 0.776194| 0.597425|
| MTOR   | 0.319847| 0.917944| 0.582309| 0.751984|
| NPY5R  | 0.439784| 0.830681| 0.832215| 0.432467|
| NR3C1  | 0.468443| 0.844652| 0.806337| 0.549943|
| NTRK1  | 0.241894| 0.932487| 0.596805| 0.522343|
| OPRD1  | 0.582086| 0.818981| 0.929093| 0.562846|
| OPRK1  | 0.46643 | 0.893228| 0.828061| 0.673175|
| OPRl1  | 0.408054| 0.86684 | 0.814934| 0.543923|
| OPRM1  | 0.557164| 0.845981| 0.829018| 0.666269|
| P2RX7  | 0.429232| 0.732701| 0.66756 | 0.405481|
| PARP1  | 0.400146| 0.836688| 0.70864 | 0.534723|
| PDE5A  | 0.476567| 0.893866| 0.797992| 0.718972|
| PDGFRB | 0.355894| 0.895651| 0.657221| 0.677594|
| PGR    | 0.419246| 0.883677| 0.777748| 0.628521|
| PIK3CA | 0.391685| 0.893135| 0.671158| 0.673145|
| PIM1   | 0.306474| 0.937288| 0.567257| 0.798435|
| PIM2   | 0.253489| 0.929945| 0.465687| 0.801314|
| PLK1   | 0.269711| 0.869752| 0.428797| 0.674745|
| PPARA  | 0.49259 | 0.771811| 0.809492| 0.43253|
| PPARD  | 0.441095| 0.854455| 0.739323| 0.601654|
| PPARG  | 0.419007| 0.850835| 0.749592| 0.549319|
| PRKACA | 0.277435| 0.874807| 0.473215| 0.645217|
| PRKCD  | 0.321108| 0.923522| 0.523234| 0.784521|
| PTGDR2 | 0.449752| 0.818296| 0.784633| 0.462896|
| PTGS2  | 0.443217| 0.827081| 0.705818| 0.595601|
| PTPN1  | 0.277734| 0.887436| 0.499561| 0.628921|
| REN    | 0.584195| 0.804456| 0.868984| 0.58534|
| ROCK1  | 0.341813| 0.890804| 0.589206| 0.674189|
| ROCK2  | 0.410589| 0.882518| 0.675734| 0.681567|
| S1PR1  | 0.478224| 0.878172| 0.924133| 0.55715|
| SCN9A  | 0.446982| 0.756159| 0.672242| 0.489453|
| SIGMAR1| 0.466162| 0.814052| 0.795478| 0.495135|
| SLC6A2 | 0.431308| 0.833322| 0.747661| 0.525409|
| SLC6A3 | 0.44193 | 0.843308| 0.690123| 0.609766|
| SRC    | 0.397001| 0.854952| 0.63397 | 0.629273|
| TACR1  | 0.500024| 0.854231| 0.837288| 0.597805|
| TRPV1  | 0.475721| 0.82854 | 0.779382| 0.558607|
| VDR    | 0.376877| 0.893166| 0.597983| 0.735373|
Detailed Results Biological Activity Prediction
Augmented model

| Gene Symbol | RMSE_train | R2_train | RMSE_test | R2_test |
|-------------|------------|----------|-----------|---------|
| ABL1        | 0.391656   | 0.827128 | 0.61209   | 0.617716|
| ACHE        | 0.483324   | 0.82706  | 0.809392  | 0.538121|
| ADAM17      | 0.435869   | 0.83693  | 0.700803  | 0.622293|
| ADORA2A     | 0.488851   | 0.808837 | 0.732699  | 0.603049|
| ADORA2B     | 0.44868    | 0.803839 | 0.653338  | 0.586217|
| ADORA3      | 0.4938     | 0.82513  | 0.769194  | 0.587006|
| ADRA1A      | 0.474645   | 0.860037 | 0.814547  | 0.564295|
| ADRA1D      | 0.46054    | 0.85318  | 0.746025  | 0.663273|
| ADRB1       | 0.441222   | 0.844254 | 0.720587  | 0.607019|
| ADRB2       | 0.427771   | 0.918238 | 0.805339  | 0.681634|
| ADRB3       | 0.46066    | 0.837282 | 0.72401   | 0.602141|
| AKT1        | 0.424124   | 0.888667 | 0.609381  | 0.786709|
| AKT2        | 0.306854   | 0.919959 | 0.433028  | 0.831577|
| ALK         | 0.362273   | 0.910441 | 0.630611  | 0.7013  |
| ALOX5       | 0.419576   | 0.718234 | 0.608661  | 0.426286|
| AR          | 0.4876     | 0.83127  | 0.872122  | 0.526351|
| AURKA       | 0.403901   | 0.848137 | 0.619312  | 0.668713|
| AURKB       | 0.383964   | 0.878276 | 0.636813  | 0.697615|
| BACE1       | 0.506486   | 0.766132 | 0.715787  | 0.511447|
| CA1         | 0.431847   | 0.831187 | 0.549425  | 0.700754|
| CA12        | 0.372581   | 0.875944 | 0.545781  | 0.76829 |
| CA2         | 0.392396   | 0.843741 | 0.585537  | 0.663244|
| CA9         | 0.402307   | 0.851308 | 0.596687  | 0.678456|
| CASP1       | 0.409233   | 0.902437 | 0.663016  | 0.718793|
| CCKBR       | 0.588082   | 0.779749 | 0.8332    | 0.600948|
| CCR2        | 0.524428   | 0.751201 | 0.803997  | 0.462313|
| Gene  | Value1  | Value2  | Value3  | Value4  |
|-------|---------|---------|---------|---------|
| CCR5  | 0.503417| 0.822251| 0.785162| 0.553117|
| CDK1  | 0.364178| 0.863288| 0.578893| 0.691975|
| CDK2  | 0.375333| 0.899998| 0.65183 | 0.679845|
| CHEK1 | 0.412467| 0.906337| 0.673386| 0.748577|
| CHRM1 | 0.472974| 0.864301| 0.893917| 0.559368|
| CHRM2 | 0.47894 | 0.866116| 0.779187| 0.644067|
| CHRM3 | 0.509352| 0.880868| 0.796349| 0.733597|
| CHRNA7| 0.545683| 0.782189| 0.877902| 0.416614|
| CLK4  | 0.404911| 0.820492| 0.650752| 0.598301|
| CNR1  | 0.499041| 0.82314 | 0.732095| 0.638208|
| CNR2  | 0.486875| 0.826751| 0.737879| 0.610197|
| CRHR1 | 0.470222| 0.773072| 0.686378| 0.501883|
| CSF1R | 0.403451| 0.896299| 0.644015| 0.73834 |
| CTSK  | 0.517701| 0.846199| 0.831942| 0.578268|
| CTSS  | 0.451625| 0.834611| 0.759972| 0.543461|
| CYP19A1| 0.49013 | 0.812804| 0.834011| 0.47432 |
| DHFR  | 0.638348| 0.828347| 1.001934| 0.562257|
| DPP4  | 0.462853| 0.789795| 0.705694| 0.559726|
| DRD1  | 0.516969| 0.791061| 0.84996 | 0.514565|
| DRD3  | 0.491006| 0.845768| 0.771762| 0.614314|
| DRD4  | 0.474372| 0.793332| 0.774309| 0.452617|
| DYRK1A| 0.391861| 0.774946| 0.593033| 0.494985|
| EDNRA | 0.553641| 0.831478| 0.829172| 0.618705|
| EGFR  | 0.471389| 0.848708| 0.747767| 0.630885|
| EPHX2 | 0.489661| 0.788611| 0.782925| 0.462102|
| ERBB2 | 0.394923| 0.835696| 0.560497| 0.683521|
| ESR1  | 0.438655| 0.874637| 0.747368| 0.661792|
| ESR2  | 0.471316| 0.82475 | 0.79314 | 0.507156|
| F10   | 0.515034| 0.861625| 0.807887| 0.681458|
| F2    | 0.522939| 0.850895| 0.795855| 0.678746|
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| FAAH   | 0.529855| 0.812364| 0.807123| 0.567572|
| FGFR1  | 0.33186 | 0.873586| 0.528079| 0.685673|
| FLT1   | 0.380861| 0.858593| 0.640509| 0.634143|
| FLT3   | 0.390533| 0.886401| 0.690762| 0.654504|
| GHSR   | 0.528352| 0.748306| 0.827759| 0.353241|
| GNRHR  | 0.651463| 0.704895| 0.779475| 0.571349|
| GRM5   | 0.493778| 0.753835| 0.743541| 0.500006|
| GSK3A  | 0.326653| 0.870288| 0.498015| 0.676276|
| GSK3B  | 0.404594| 0.853648| 0.640287| 0.626728|
| HDAC1  | 0.409744| 0.829324| 0.68898 | 0.544812|
| HPGD   | 0.328278| 0.53038 | 0.426883| 0.372395|
| HRH3   | 0.486531| 0.79568 | 0.696746| 0.60006 |
| HSD11B1| 0.529623| 0.719811| 0.758364| 0.482519|
| HSP90AA1|0.332854| 0.881642| 0.529756| 0.695969|
| HTR2A  | 0.468105| 0.847087| 0.772319| 0.586357|
| HTR2C  | 0.472821| 0.78781 | 0.733308| 0.489671|
| HTR6   | 0.461801| 0.833969| 0.773748| 0.532445|
| HTR7   | 0.472384| 0.785086| 0.765704| 0.481417|
| IGF1R  | 0.345364| 0.873992| 0.445232| 0.793183|
| INSR   | 0.298785| 0.85626 | 0.367613| 0.748355|
| ITK    | 0.332881| 0.902446| 0.540686| 0.69657 |
| JAK2   | 0.376016| 0.901693| 0.657719| 0.690164|
| JAK3   | 0.369665| 0.833417| 0.557399| 0.595265|
| KCNH2  | 0.375089| 0.778064| 0.639404| 0.404628|
| KDR    | 0.429687| 0.855481| 0.710769| 0.638369|
| KIT    | 0.381266| 0.866972| 0.67155 | 0.580183|
| LCK    | 0.386504| 0.888735| 0.643991| 0.685003|
| MAOB   | 0.493586| 0.80015 | 0.857941| 0.408251|
| MAPK14 | 0.43662 | 0.845714| 0.705503| 0.590673|
| MAPK8  | 0.323676| 0.865468| 0.532037| 0.690627|
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| MAPK9  | 0.313757| 0.863519| 0.482046| 0.671265|
| MAPKAPK2| 0.335893| 0.895211| 0.466454| 0.79768 |
| MC4R   | 0.543342| 0.75523 | 0.755209| 0.539072|
| MCHR1  | 0.479404| 0.772348| 0.737637| 0.509469|
| MET    | 0.374524| 0.887642| 0.606851| 0.73225 |
| MMP1   | 0.434695| 0.830342| 0.646003| 0.614351|
| MMP13  | 0.471431| 0.844127| 0.759325| 0.598611|
| MMP2   | 0.452296| 0.886602| 0.801673| 0.648447|
| MMP3   | 0.445396| 0.822075| 0.731301| 0.556404|
| MMP9   | 0.443492| 0.861442| 0.739456| 0.627294|
| MTOR   | 0.46535 | 0.880139| 0.768625| 0.710081|
| NPY5R  | 0.48765 | 0.791832| 0.81342 | 0.440046|
| NR3C1  | 0.509845| 0.815813| 0.780582| 0.572355|
| NTRK1  | 0.304811| 0.89318 | 0.562126| 0.578924|
| OPRD1  | 0.624437| 0.791651| 0.906759| 0.577935|
| OPRL1  | 0.514682| 0.869959| 0.81698 | 0.678388|
| OPRM1  | 0.464951| 0.827208| 0.830186| 0.505011|
| P2RX7  | 0.451454| 0.703923| 0.661416| 0.412124|
| PARP1  | 0.44889 | 0.794508| 0.705237| 0.538537|
| PDE5A  | 0.487228| 0.888993| 0.843174| 0.685396|
| PDGFRB | 0.415556| 0.85786 | 0.699975| 0.637222|
| PGR    | 0.46236 | 0.858356| 0.79069 | 0.612464|
| PIK3CA | 0.442589| 0.863451| 0.66311 | 0.67768 |
| PIM1   | 0.339   | 0.923589| 0.587345| 0.786273|
| PIM2   | 0.292331| 0.906833| 0.471866| 0.797908|
| PLK1   | 0.301631| 0.837054| 0.443016| 0.656059|
| PPARA  | 0.529993| 0.735816| 0.80248 | 0.427  |
| PPARD  | 0.501673| 0.811603| 0.741052| 0.59705 |
| PPARG  | 0.454712| 0.824647| 0.703102| 0.592409|
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| PRKACA | 0.32414 | 0.82781 | 0.5084  | 0.587839|
| PRKCD  | 0.358083| 0.904506| 0.530458| 0.779047|
| PTGDR2 | 0.508535| 0.767521| 0.757318| 0.493816|
| PTGS2  | 0.500585| 0.779255| 0.696415| 0.607257|
| PTPN1  | 0.317422| 0.853399| 0.486506| 0.636844|
| REN    | 0.621654| 0.778324| 0.887979| 0.568382|
| ROCK1  | 0.381253| 0.864896| 0.580158| 0.677347|
| ROCK2  | 0.431194| 0.870751| 0.669503| 0.677877|
| S1PR1  | 0.520008| 0.855891| 0.846751| 0.61829 |
| SCN9A  | 0.453935| 0.748702| 0.67755 | 0.482101|
| SIGMAR1| 0.498822| 0.78687 | 0.808203| 0.481596|
| SLC6A2 | 0.473616| 0.799059| 0.740727| 0.528622|
| SLC6A3 | 0.475889| 0.818495| 0.680153| 0.619581|
| SRC    | 0.436091| 0.825007| 0.621299| 0.637691|
| TACR1  | 0.545989| 0.826357| 0.771087| 0.651392|
| TRPV1  | 0.509238| 0.803496| 0.777344| 0.557124|
| VDR    | 0.42594 | 0.863476| 0.685624| 0.659814|
Detailed results on bioactivities using the Combined model

| Gene_Symbol | RMSE_train | R2_train | RMSE_test | R2_test |
|-------------|------------|----------|-----------|---------|
| ABL1        | 0.206298   | 0.951973 | 0.558613  | 0.678433 |
| ACHE        | 0.316561   | 0.925781 | 0.777284  | 0.584371 |
| ADAM17      | 0.248802   | 0.946877 | 0.663541  | 0.663298 |
| ADORA2A     | 0.286617   | 0.93427  | 0.694326  | 0.64388 |
| ADORA2B     | 0.260213   | 0.933873 | 0.636431  | 0.618746 |
| ADORA3      | 0.326777   | 0.923422 | 0.779192  | 0.584347 |
| ADRA1A      | 0.310485   | 0.940134 | 0.781936  | 0.605527 |
| ADRA1D      | 0.332591   | 0.944284 | 0.705648  | 0.623655 |
| ADRB1       | 0.282417   | 0.923517 | 0.680911  | 0.715772 |
| ADRB2       | 0.280009   | 0.96498  | 0.772141  | 0.56036 |
| ADRB3       | 0.269552   | 0.944284 | 0.705648  | 0.623655 |
| AKT1        | 0.251131   | 0.960963 | 0.632493  | 0.774043 |
| AKT2        | 0.207717   | 0.963596 | 0.380975  | 0.862947 |
| ALK         | 0.209744   | 0.969985 | 0.591488  | 0.734167 |
| ALOX5       | 0.272735   | 0.880912 | 0.567513  | 0.498686 |
| AR          | 0.326987   | 0.924137 | 0.873306  | 0.530971 |
| AURKA       | 0.221903   | 0.954154 | 0.576972  | 0.712729 |
| AURKB       | 0.213774   | 0.962282 | 0.604271  | 0.724682 |
| BACE1       | 0.309571   | 0.91263  | 0.699941  | 0.542096 |
| CA1         | 0.241656   | 0.947097 | 0.551373  | 0.703278 |
| CA12        | 0.214921   | 0.958886 | 0.552578  | 0.763067 |
| CA2         | 0.230419   | 0.946253 | 0.518601  | 0.737193 |
| CA9         | 0.225505   | 0.953463 | 0.54828   | 0.726171 |
| CASP1       | 0.235935   | 0.967618 | 0.652108  | 0.731127 |
| CCKBR       | 0.434981   | 0.87972  | 0.789642  | 0.643014 |
| CCR2        | 0.334621   | 0.898654 | 0.794792  | 0.48265 |
| CCR5        | 0.299886   | 0.936982 | 0.756141  | 0.592012 |
| CDK1        | 0.199109   | 0.959126 | 0.512894  | 0.755824 |
| CDK2        | 0.208584   | 0.96607  | 0.676458  | 0.662738 |
| CHEK1       | 0.226676   | 0.971709 | 0.640112  | 0.771291 |
| CHRM1       | 0.300844   | 0.945086 | 0.883212  | 0.57378 |
| CHRM2       | 0.31254    | 0.942901 | 0.765526  | 0.66241 |
| CHRM3       | 0.317695   | 0.953725 | 0.706711  | 0.791522 |
| CHRNA7      | 0.39232    | 0.887443 | 0.930657  | 0.387507 |
| CLK4        | 0.225163   | 0.944504 | 0.630395  | 0.619228 |
| CNR1        | 0.329539   | 0.923102 | 0.718551  | 0.658166 |
| CNR2        | 0.29941    | 0.934475 | 0.714849  | 0.639133 |
| CRHR1       | 0.268734   | 0.925879 | 0.653135  | 0.554787 |
| CSF1R       | 0.265925   | 0.954847 | 0.640827  | 0.740349 |
| CTSK        | 0.359195   | 0.925969 | 0.817168  | 0.597735 |
| CTS5        | 0.267147   | 0.942108 | 0.656775  | 0.66083 |
| CYP19A1     | 0.271379   | 0.942606 | 0.842616  | 0.476361 |
| DHFR        | 0.409279   | 0.929379 | 1.01431   | 0.567761 |
| Gene   | Value1 | Value2 | Value3 | Value4 |
|--------|--------|--------|--------|--------|
| DPP4   | 0.285293 | 0.920115 | 0.653857 | 0.623939 |
| DRD1   | 0.350931 | 0.903617 | 0.925557 | 0.446277 |
| DRD3   | 0.313099 | 0.937257 | 0.770997 | 0.615358 |
| DRD4   | 0.287455 | 0.924251 | 0.761317 | 0.478539 |
| DYRK1A | 0.226334 | 0.924935 | 0.551734 | 0.561336 |
| EDNRA  | 0.343053 | 0.935276 | 0.810721 | 0.635912 |
| EGFR   | 0.264955 | 0.952192 | 0.740926 | 0.643192 |
| EPHX2  | 0.342906 | 0.896387 | 0.744952 | 0.510932 |
| ERBB2  | 0.234829 | 0.916310 | 0.538777 | 0.708806 |
| ESR1   | 0.254966 | 0.957645 | 0.749739 | 0.659670 |
| ESR2   | 0.284116 | 0.936326 | 0.810858 | 0.495771 |
| F10    | 0.312452 | 0.957932 | 0.522582 | 0.696779 |
| F2     | 0.241116 | 0.936326 | 0.810858 | 0.495771 |
| FAAH   | 0.354098 | 0.916310 | 0.574246 | 0.574246 |
| FGFR1  | 0.191828 | 0.957932 | 0.522582 | 0.696779 |
| FLT1   | 0.214447 | 0.955159 | 0.563869 | 0.702749 |
| FLT3   | 0.223692 | 0.962735 | 0.671025 | 0.672371 |
| GHSR   | 0.340175 | 0.895487 | 0.756875 | 0.443094 |
| GNRHR  | 0.495379 | 0.829265 | 0.837442 | 0.527127 |
| GRM5   | 0.272717 | 0.924856 | 0.72775  | 0.527981 |
| GSK3A  | 0.182533 | 0.959592 | 0.454632 | 0.754102 |
| GSK3B  | 0.2254  | 0.954557 | 0.61449  | 0.657639 |
| HDAC1  | 0.241828 | 0.940478 | 0.626598 | 0.619149 |
| HPGD   | 0.198257 | 0.829901 | 0.445062 | 0.356863 |
| HRH3   | 0.362574 | 0.886594 | 0.72023  | 0.585162 |
| HSD11B1| 0.400956 | 0.839408 | 0.782626 | 0.456629 |
| HSP90AA1| 0.186326 | 0.962956 | 0.509832 | 0.721266 |
| HTR2A  | 0.290569 | 0.941008 | 0.735598 | 0.63213 |
| HTR2C  | 0.298729 | 0.915326 | 0.688168 | 0.546307 |
| HTR6   | 0.297955 | 0.930789 | 0.711453 | 0.597754 |
| HTR7   | 0.248617 | 0.940381 | 0.719072 | 0.536167 |
| IGF1R  | 0.208017 | 0.954237 | 0.402755 | 0.83075 |
| INSR   | 0.17537  | 0.950661 | 0.367473 | 0.755836 |
| ITK    | 0.193004 | 0.967163 | 0.565616 | 0.675914 |
| JAK2   | 0.234838 | 0.961607 | 0.637093 | 0.707538 |
| JAK3   | 0.213668 | 0.944366 | 0.506395 | 0.657579 |
| KCNH2  | 0.22123  | 0.922731 | 0.596583 | 0.473338 |
| KDR    | 0.241795 | 0.954263 | 0.652572 | 0.693961 |
| KIT    | 0.191701 | 0.966391 | 0.648837 | 0.603106 |
| LCK    | 0.214887 | 0.965591 | 0.60336 | 0.722575 |
| MAOB   | 0.285953 | 0.932892 | 0.834083 | 0.441061 |
| MAPK14 | 0.258373 | 0.94596  | 0.685918 | 0.617252 |
| MAPK8  | 0.186508 | 0.955266 | 0.508969 | 0.71753 |
| MAPK9  | 0.180443 | 0.95492  | 0.424372 | 0.747357 |
| MAPKAPK2| 0.199052 | 0.963508 | 0.46819  | 0.808204 |
| MC4R   | 0.379861 | 0.880343 | 0.741193 | 0.560735 |
| Gene   | Value1   | Value2   | Value3   | Value4   |
|--------|----------|----------|----------|----------|
| MCHR1  | 0.283749 | 0.920382 | 0.713824 | 0.546953 |
| MET    | 0.211169 | 0.964236 | 0.577767 | 0.755813 |
| MMP1   | 0.256735 | 0.940876 | 0.620326 | 0.646093 |
| MMP13  | 0.286161 | 0.942606 | 0.723201 | 0.643229 |
| MMP2   | 0.269886 | 0.959586 | 0.776835 | 0.674197 |
| MMP3   | 0.26641  | 0.936416 | 0.670422 | 0.623077 |
| MMP9   | 0.263269 | 0.951188 | 0.741729 | 0.632901 |
| MTOR   | 0.229953 | 0.970798 | 0.711027 | 0.748265 |
| NPY5R  | 0.261002 | 0.940547 | 0.810981 | 0.467157 |
| NR3C1  | 0.335298 | 0.920351 | 0.750368 | 0.607957 |
| NTRK1  | 0.159021 | 0.970872 | 0.560701 | 0.583291 |
| OPRD1  | 0.460809 | 0.886492 | 0.90379  | 0.589825 |
| OPRK1  | 0.326723 | 0.947593 | 0.805359 | 0.690791 |
| OPRL1  | 0.287514 | 0.933916 | 0.783634 | 0.549366 |
| OPRM1  | 0.435529 | 0.9058   | 0.81616  | 0.679574 |
| P2RX7  | 0.304949 | 0.865005 | 0.636871 | 0.463937 |
| PARP1  | 0.282813 | 0.918375 | 0.672585 | 0.579111 |
| PDE5A  | 0.341468 | 0.945476 | 0.84484  | 0.69215  |
| PDGFRB | 0.21918  | 0.960564 | 0.633898 | 0.701701 |
| PGR    | 0.273476 | 0.950437 | 0.765383 | 0.642248 |
| PIK3CA | 0.234429 | 0.961695 | 0.668303 | 0.672411 |
| PIM1   | 0.21632  | 0.969016 | 0.555629 | 0.807533 |
| PIM2   | 0.172049 | 0.967727 | 0.449478 | 0.815307 |
| PLK1   | 0.194322 | 0.932562 | 0.434646 | 0.67264  |
| PPARA  | 0.351923 | 0.883551 | 0.839983 | 0.406571 |
| PPARD  | 0.291829 | 0.936277 | 0.713224 | 0.627577 |
| PPARG  | 0.273049 | 0.936742 | 0.729574 | 0.578892 |
| PRKACA | 0.190592 | 0.940701 | 0.471264 | 0.643625 |
| PRKCD  | 0.215146 | 0.965565 | 0.543846 | 0.771233 |
| PTGDR2 | 0.28127  | 0.929023 | 0.757473 | 0.501684 |
| PTGS2  | 0.288023 | 0.9269   | 0.709553 | 0.607711 |
| PTPN1  | 0.203241 | 0.940099 | 0.469528 | 0.663897 |
| REN    | 0.427246 | 0.895504 | 0.867667 | 0.590599 |
| ROCK1  | 0.215101 | 0.957403 | 0.5255   | 0.731566 |
| ROCK2  | 0.275904 | 0.946964 | 0.659185 | 0.693623 |
| S1PR1  | 0.308967 | 0.949132 | 0.877469 | 0.593745 |
| SCN9A  | 0.264152 | 0.914763 | 0.647838 | 0.531714 |
| SIGMAR1| 0.340304 | 0.900903 | 0.804496 | 0.500548 |
| SLC6A2 | 0.304194 | 0.917095 | 0.735474 | 0.540423 |
| SLC6A3 | 0.300728 | 0.927449 | 0.648582 | 0.656814 |
| SRC    | 0.273111 | 0.931319 | 0.627292 | 0.634998 |
| TACR1  | 0.353287 | 0.927327 | 0.834879 | 0.604734 |
| TRPV1  | 0.315027 | 0.924834 | 0.800165 | 0.551962 |
| VDR    | 0.243347 | 0.955575 | 0.649073 | 0.692942 |
**Detailed results of bioactivities using the SVR model**

| Gene_Symbol | RMSE_train | RMSE_test | R2_train | R2_test |
|-------------|------------|-----------|----------|---------|
| ABL1        | 0.185194   | 0.614615  | 0.965582 | 0.606576|
| ACH    | 0.32912    | 0.70019   | 0.919713 | 0.644237|
| ADAM17      | 0.169504   | 0.594161  | 0.976406 | 0.725887|
| ADORA2A     | 0.296635   | 0.658876  | 0.933689 | 0.675812|
| ADORA2B     | 0.186061   | 0.61007   | 0.966699 | 0.637764|
| ADORA3       | 0.328443   | 0.700011  | 0.924065 | 0.651177|
| ADRA1A       | 0.264756   | 0.747027  | 0.955849 | 0.622354|
| ADRA1D       | 0.29994    | 0.71299   | 0.939938 | 0.696083|
| ADRB1        | 0.282635   | 0.677339  | 0.937379 | 0.642848|
| ADRB2        | 0.286934   | 0.641331  | 0.978336 | 0.794666|
| ADRB3        | 0.245891   | 0.577861  | 0.93146  | 0.696141|
| AKT1         | 0.188628   | 0.596658  | 0.978336 | 0.794666|
| AKT2         | 0.107589   | 0.410271  | 0.990086 | 0.842437|
| ALK          | 0.116641   | 0.577861  | 0.991014 | 0.743354|
| ALOX5        | 0.266698   | 0.576248  | 0.89144  | 0.469967|
| AR           | 0.322655   | 0.826243  | 0.92734  | 0.556002|
| AURKA        | 0.130168   | 0.580075  | 0.984633 | 0.703645|
| AURKB        | 0.177342   | 0.621831  | 0.977881 | 0.706842|
| BACE1        | 0.289113   | 0.617412  | 0.926773 | 0.618453|
| CA1          | 0.222987   | 0.526274  | 0.955566 | 0.724183|
| CA12         | 0.087038   | 0.536323  | 0.993472 | 0.777701|
| CA2          | 0.171591   | 0.522412  | 0.970347 | 0.727941|
| CA9          | 0.149584   | 0.541832  | 0.979115 | 0.727654|
| CASP1        | 0.205572   | 0.554131  | 0.97603  | 0.800695|
| CCKBR        | 0.422571   | 0.731618  | 0.888963 | 0.687612|
| CCR2         | 0.353643   | 0.666351  | 0.894566 | 0.613838|
| CCR5         | 0.294916   | 0.635023  | 0.941513 | 0.695608|
| CDK1         | 0.066722   | 0.573859  | 0.995456 | 0.694234|
| CDK2         | 0.171489   | 0.641331  | 0.97878  | 0.688573|
| CHEK1        | 0.124957   | 0.62042   | 0.991438 | 0.782421|
| CHRM1        | 0.338907   | 0.800901  | 0.936682 | 0.630765|
| CHRM2        | 0.289674   | 0.760078  | 0.950967 | 0.65552|
| CHRM3        | 0.354411   | 0.743165  | 0.94267  | 0.766108|
| CHRNA7       | 0.441708   | 0.815506  | 0.860721 | 0.478982|
| CLK4         | 0.190357   | 0.67177   | 0.969815 | 0.560984|
| CNR1         | 0.303116   | 0.642851  | 0.934641 | 0.717171|
| CNR2         | 0.283189   | 0.6693    | 0.942478 | 0.675429|
| CRHR1        | 0.164358   | 0.596783  | 0.972903 | 0.612404|
| CSF1R        | 0.256591   | 0.649347  | 0.95952  | 0.726185|
| CTSK         | 0.401049   | 0.775466  | 0.910533 | 0.627301|
| CTSS         | 0.209439   | 0.590364  | 0.966696 | 0.719058|
| CYP19A1      | 0.304845   | 0.749211  | 0.933677 | 0.571726|
| DHFR         | 0.416729   | 0.903808  | 0.926814 | 0.628217|
| Gene   | Value1  | Value2  | Value3  | Value4  |
|--------|---------|---------|---------|---------|
| DPP4   | 0.228322| 0.609309| 0.948622| 0.668076|
| DRD1   | 0.389004| 0.750376| 0.885103| 0.617612|
| DRD3   | 0.295467| 0.681352| 0.945445| 0.617612|
| DRD4   | 0.236273| 0.668839| 0.951542| 0.571021|
| Dyrk1a | 0.177273| 0.558207| 0.959274| 0.531306|
| Ednra  | 0.310926| 0.690718| 0.947669| 0.722196|
| Egfr   | 0.197623| 0.668461| 0.973812| 0.700161|
| Ephi2  | 0.347541| 0.748645| 0.896322| 0.491367|
| Erbb2  | 0.229236| 0.509764| 0.949947| 0.735971|
| ESR1   | 0.194849| 0.723354| 0.976464| 0.675748|
| ESR2   | 0.33452 | 0.763538| 0.925095| 0.52538 |
| F10    | 0.289859| 0.67794 | 0.950737| 0.773409|
| F2     | 0.248381| 0.666103| 0.966693| 0.770077|
| FaaH   | 0.382898| 0.736538| 0.905143| 0.623748|
| Fgfr1  | 0.114259| 0.520928| 0.985067| 0.688712|
| Flt1   | 0.15133 | 0.63979 | 0.983564| 0.657296|
| Flt3   | 0.171594| 0.714967| 0.978184| 0.61917 |
| Ghsr   | 0.292276| 0.698092| 0.921678| 0.516784|
| Gnhr   | 0.573733| 0.762312| 0.778397| 0.583027|
| Grm5   | 0.178541| 0.627948| 0.969489| 0.630698|
| Gsk3a  | 0.116874| 0.506394| 0.986613| 0.595658|
| Gsk3b  | 0.137329| 0.606966| 0.983564| 0.657296|
| Hdac1  | 0.325472| 0.55667 | 0.894884| 0.686902|
| Hpgd   | 0.120303| 0.400431| 0.953426| 0.444259|
| Hrh3   | 0.355461| 0.680458| 0.894478| 0.614622|
| Hsd11b1| 0.458122| 0.725424| 0.796125| 0.510449|
| Hsp90aa1| 0.060567| 0.46122 | 0.996582| 0.765892|
| Htr2a  | 0.285355| 0.7058  | 0.945002| 0.645554|
| Htr2c  | 0.253909| 0.653496| 0.942707| 0.572908|
| Htr6   | 0.214436| 0.642899| 0.965858| 0.657359|
| Htr7   | 0.194823| 0.657101| 0.965991| 0.602549|
| Igf1r  | 0.162801| 0.388727| 0.973049| 0.837972|
| Insr   | 0.065803| 0.455177| 0.992823| 0.619515|
| Itk    | 0.110813| 0.518228| 0.989006| 0.717493|
| Jak2   | 0.159794| 0.666446| 0.981923| 0.68051 |
| Jak3   | 0.114836| 0.543641| 0.983793| 0.598501|
| Kcnh2  | 0.138921| 0.555443| 0.972641| 0.520159|
| Kdr    | 0.166137| 0.663386| 0.978846| 0.680618|
| Kit    | 0.139398| 0.60384 | 0.98209 | 0.62916 |
| Lck    | 0.170436| 0.646383| 0.980709| 0.681596|
| Maob   | 0.320031| 0.768364| 0.920256| 0.489306|
| Mapk14 | 0.163466| 0.623158| 0.979377| 0.669392|
| Mapk8  | 0.099239| 0.505868| 0.987451| 0.718817|
| Mapk9  | 0.080408| 0.441814| 0.991318| 0.721171|
| Mapkapk2| 0.091687| 0.45494 | 0.992291| 0.807092|
| Mc4r   | 0.354664| 0.633084| 0.896669| 0.667435|
| Gene   | MCHR1   | 0.215083 | 0.642111 | 0.955827 | 0.61969 |
|--------|---------|----------|----------|----------|---------|
| MET    | 0.13651 | 0.538442 | 0.985221 | 0.784688 |
| MMP1   | 0.242328 | 0.593555 | 0.946803 | 0.662139 |
| MMP13  | 0.328806 | 0.695963 | 0.923949 | 0.647964 |
| MMP2   | 0.246612 | 0.755626 | 0.966628 | 0.68003 |
| MMP3   | 0.263269 | 0.65841  | 0.939913 | 0.633534 |
| MMP9   | 0.266111 | 0.740258 | 0.951859 | 0.620088 |
| MTKOR  | 0.204005 | 0.690035 | 0.979872 | 0.634792 |
| NR3C1  | 0.309608 | 0.705999 | 0.932922 | 0.645299 |
| NTRK1  | 0.096557 | 0.523373 | 0.989501 | 0.614797 |
| OPRD1  | 0.480979 | 0.833996 | 0.973211 | 0.636529 |
| OPRK1  | 0.36086  | 0.741872 | 0.938574 | 0.730783 |
| OPRL1  | 0.257543 | 0.722106 | 0.94624  | 0.611553 |
| OPRM1  | 0.453105 | 0.732095 | 0.899494 | 0.736603 |
| P2RX7  | 0.259605 | 0.650234 | 0.904628 | 0.430616 |
| PARP1  | 0.239059 | 0.625782 | 0.945356 | 0.625496 |
| PDE5A  | 0.313263 | 0.771852 | 0.953383 | 0.733649 |
| PDGFRB | 0.137501 | 0.697576 | 0.984706 | 0.634792 |
| PGR    | 0.128689 | 0.639331 | 0.989092 | 0.740203 |
| PIK3CA | 0.164467 | 0.562775 | 0.982204 | 0.762527 |
| PIM1   | 0.073638 | 0.576178 | 0.996527 | 0.792539 |
| PIM2   | 0.053638 | 0.474807 | 0.996897 | 0.79059 |
| PLK1   | 0.150522 | 0.395537 | 0.958858 | 0.718815 |
| PPARA  | 0.328187 | 0.780685 | 0.905939 | 0.4667 |
| PPARD  | 0.307733 | 0.643645 | 0.936092 | 0.687488 |
| PPARG  | 0.22617 | 0.65974  | 0.958974 | 0.636192 |
| PRKACA | 0.162015 | 0.470888 | 0.961092 | 0.638177 |
| PRKCD  | 0.120404 | 0.512236 | 0.989689 | 0.793034 |
| PTGFR2 | 0.213569 | 0.66181 | 0.959981 | 0.598447 |
| PTGS2  | 0.278587 | 0.63852  | 0.951857 | 0.662132 |
| PTNP1  | 0.049108 | 0.421377 | 0.996623 | 0.711722 |
| REN    | 0.425341 | 0.726459 | 0.899341 | 0.704993 |
| ROCK1  | 0.158043 | 0.594678 | 0.97827 | 0.658431 |
| ROCK2  | 0.242137 | 0.655294 | 0.959398 | 0.68937 |
| S1PR1  | 0.267412 | 0.835533 | 0.962232 | 0.621598 |
| SCN9A  | 0.179124 | 0.551783 | 0.964353 | 0.644534 |
| SIGMAR1| 0.346971 | 0.747507 | 0.899193 | 0.538575 |
| SLC6A2 | 0.336916 | 0.706539 | 0.900801 | 0.559941 |
| SLC6A3 | 0.315193 | 0.62562 | 0.921698 | 0.672803 |
| SRC    | 0.244284 | 0.572583 | 0.946771 | 0.687487 |
| TACR1  | 0.362814 | 0.719143 | 0.925448 | 0.692533 |
| TRPV1  | 0.294321 | 0.686775 | 0.934833 | 0.647703 |
| VDR    | 0.24332 | 0.616355 | 0.958015 | 0.701363 |
Supplementary Table: Performance comparison of Transformer models and SVR. The RMSE and $R^2$ of the models trained on the same 20 random train/test splits, the mean plus/minus the standard deviation are shown for each of the three physical chemistry properties. The best performance is denoted with bold.

| Property    | EncRegr | EncRegrTL_spanZ | EncRegrTL_augZ | EncRegrTL_bothZ | SVR       |
|-------------|---------|-----------------|----------------|-----------------|-----------|
| Lipophilicity |        |                 |                |                 |           |
| Train RMSE  | 0.729 ± 0.015 | 0.234 ± 0.005 | 0.242 ± 0.005 | 0.164 ± 0.004 | 0.037 ± 0.005 |
| Test RMSE   | 0.946 ± 0.021 | 0.621 ± 0.021 | 0.618 ± 0.023 | **0.598 ± 0.021** | 0.746 ± 0.017 |
| Train $R^2$ | 0.634 ± 0.014 | 0.962 ± 0.001 | 0.961 ± 0.002 | 0.982 ± 0.001 | 0.999 ± 0.000 |
| Test $R^2$  | 0.398 ± 0.022 | 0.736 ± 0.018 | 0.738 ± 0.018 | **0.754 ± 0.013** | 0.617 ± 0.018 |
| ESOL        |         |                 |                |                 |           |
| Train RMSE  | 0.592 ± 0.017 | 0.305 ± 0.009 | 0.315 ± 0.008 | 0.248 ± 0.009 | 0.307 ± 0.031 |
| Test RMSE   | 0.803 ± 0.063 | 0.657 ± 0.040 | 0.652 ± 0.047 | **0.633 ± 0.039** | 1.031 ± 0.084 |
| Train $R^2$ | 0.919 ± 0.004 | 0.978 ± 0.001 | 0.977 ± 0.001 | 0.986 ± 0.001 | 0.978 ± 0.004 |
| Test $R^2$  | 0.855 ± 0.024 | 0.903 ± 0.014 | 0.904 ± 0.017 | **0.910 ± 0.014** | 0.766 ± 0.044 |
| FreeSolvation |         |                 |                |                 |           |
| Train RMSE  | 1.148 ± 0.044 | 0.498 ± 0.019 | 0.525 ± 0.014 | 0.414 ± 0.018 | 0.057 ± 0.003 |
| Test RMSE   | 1.886 ± 0.278 | 1.366 ± 0.250 | 1.287 ± 0.273 | **1.230 ± 0.298** | 2.108 ± 0.570 |
| Train $R^2$ | 0.905 ± 0.010 | 0.982 ± 0.002 | 0.980 ± 0.001 | 0.988 ± 0.001 | 0.999 ± 0.000 |
| Test $R^2$  | 0.786 ± 0.067 | 0.889 ± 0.035 | 0.901 ± 0.036 | **0.908 ± 0.041** | 0.754 ± 0.092 |
Supplementary Table: Hyperparameters found for the SVR models

| Gene_Symbol | #tr_obs | C   | gamma | epsilon | test_obs | model |
|-------------|---------|-----|-------|---------|----------|-------|
| ABL1        | 1365    | 5.802512 | 0.858588 | 0.143336 | 493      | SVR   |
| ACHB        | 2313    | 5.235646 | 1     | 0.09424 | 838      | SVR   |
| ADAM17      | 1047    | 5.111516 | 0.738218 | 0.006229 | 324      | SVR   |
| ADORA2A     | 3903    | 5.135936 | 1.00E-05 | 0.186929 | 1366     | SVR   |
| ADORA2B     | 1581    | 8.171639 | 0.068761 | 0.001    | 491      | SVR   |
| ADORA3      | 2853    | 5.712518 | 0.981636 | 0.061718 | 957      | SVR   |
| ADRA1A      | 1314    | 7.330678 | 1     | 0.001   | 465      | SVR   |
| ADRA1D      | 1054    | 45.16893 | 1.00E-05 | 0.213605 | 404      | SVR   |
| ADRB1       | 966     | 5.391873 | 1     | 0.154251 | 331      | SVR   |
| ADRB3       | 1461    | 4.725941 | 1     | 0.139392 | 479      | SVR   |
| ADRB3       | 1074    | 5.7268   | 0.099707 | 0.001    | 381      | SVR   |
| AKT1        | 1677    | 5.143698 | 1.00E-05 | 0.001    | 552      | SVR   |
| AKT2        | 1027    | 5.728409 | 1     | 0.001   | 341      | SVR   |
| ALK         | 1058    | 8.742291 | 0.4629 | 0.099466 | 376      | SVR   |
| ALOX5       | 2167    | 5.620258 | 1     | 0.124465 | 724      | SVR   |
| AR          | 1954    | 7.184241 | 1     | 0.077524 | 617      | SVR   |
| AURKA       | 2062    | 5.626241 | 0.994243 | 0.046747 | 671      | SVR   |
| AURKB       | 1624    | 5.612974 | 1     | 0.14825  | 575      | SVR   |
| BACE1       | 2227    | 5.365991 | 1     | 0.147455 | 724      | SVR   |
| CA1         | 2141    | 5.567289 | 0.847257 | 0.029508 | 717      | SVR   |
| CA12        | 960     | 6.427096 | 1     | 0.001   | 339      | SVR   |
| CA2         | 2756    | 5.710749 | 0.414136 | 0.004364 | 896      | SVR   |
| CA9         | 1309    | 6.108501 | 1.00E-05 | 0.001    | 435      | SVR   |
| CASP1       | 1043    | 5.082819 | 0.507596 | 0.142454 | 335      | SVR   |
| CCKBR       | 1204    | 4.777796 | 0.852884 | 0.172617 | 407      | SVR   |
| CCR2        | 1267    | 5.010372 | 1.00E-05 | 0.251409 | 422      | SVR   |
| CCR5        | 1531    | 5.355618 | 0.993217 | 0.106631 | 485      | SVR   |
| CDK1        | 1175    | 8.456855 | 0.3824 | 0.001    | 418      | SVR   |
| CDK2        | 1548    | 5.985689 | 0.975564 | 0.095688 | 540      | SVR   |
| CHEK1       | 2029    | 6.97949  | 0.533681 | 0.001    | 689      | SVR   |
| CHRM1       | 2051    | 5.227561 | 1.00E-05 | 0.210307 | 706      | SVR   |
| CHRM2       | 1331    | 5.804398 | 1     | 0.029462 | 456      | SVR   |
| CHRM3       | 1255    | 4.232168 | 1     | 0.001   | 417      | SVR   |
| CHRNA7      | 1074    | 4.418988 | 0.948666 | 0.001    | 369      | SVR   |
| CLK4        | 1124    | 5.490684 | 0.071437 | 0.186867 | 368      | SVR   |
| CNR1        | 4010    | 5.312096 | 1     | 0.047831 | 1269     | SVR   |
| CNR2        | 3350    | 5.044598 | 1     | 0.067373 | 1126     | SVR   |
| CRHR1       | 1456    | 5.415964 | 1.00E-05 | 0.001    | 543      | SVR   |
| CSF1R       | 1488    | 5.271057 | 1     | 0.124997 | 498      | SVR   |
| CTSK        | 1214    | 5.050977 | 1.00E-05 | 0.189517 | 389      | SVR   |
| CTSS        | 1166    | 5.195509 | 1     | 0.107482 | 404      | SVR   |
| CYP19A1     | 1229    | 6.207021 | 1     | 0.185594 | 393      | SVR   |
| DHFR        | 1114    | 8.742291 | 0.4629 | 0.099466 | 402      | SVR   |
| Gene   | Value1 | Value2 | Value3 | p-value | SVR  |
|--------|--------|--------|--------|---------|------|
| DPP4   | 2480   | 5.741587 | 1.00E-05 | 0.063335 | 813  |
| DRD1   | 1307   | 5.029413 | 1       | 0.233249 | 425  |
| DRD3   | 2637   | 5.760533 | 0.618961 | 0.108905 | 854  |
| DRD4   | 1470   | 5.618841 | 1.00E-05 | 0.06868  | 486  |
| DyrK1A | 1915   | 7.13613  | 0.414324 | 0.134191 | 653  |
| EDNRA  | 1556   | 5.036787 | 0.076803 | 0.08119  | 529  |
| EGFR   | 3248   | 8.742291 | 0.4629  | 0.099466 | 1066 |
| EPHX2  | 1338   | 4.778157 | 1.00E-05 | 0.039619 | 454  |
| ERBB2  | 1479   | 5.201756 | 1       | 0.228193 | 445  |
| ESR1   | 1951   | 5.505274 | 1       | 0.052016 | 640  |
| ESR2   | 1467   | 5.12859  | 1.00E-05 | 0.304988 | 479  |
| F10    | 3532   | 5.085705 | 1.00E-05 | 0.104156 | 1105 |
| F2     | 2958   | 6.993556 | 1       | 0.058587 | 938  |
| FAAH   | 1702   | 4.962366 | 1.00E-05 | 0.20132  | 565  |
| FGFR1  | 1350   | 4.062837 | 1       | 0.016634 | 457  |
| FLT1   | 1247   | 5.296188 | 0.907542 | 0.065004 | 396  |
| FLT3   | 1555   | 5.660203 | 1       | 0.001    | 445  |
| GHSR   | 923    | 4.394825 | 1       | 0.001    | 318  |
| GNRHR  | 1276   | 4.267755 | 0.750858 | 0.386028 | 444  |
| GRM5   | 1877   | 5.832078 | 1.00E-05 | 0.096016 | 657  |
| GSK3A  | 1224   | 20.51896 | 1       | 0.103245 | 382  |
| GSK3B  | 2474   | 5.482835 | 1.00E-05 | 0.01385  | 856  |
| HDAC1  | 1459   | 1.752279 | 1       | 0.001    | 477  |
| HPGD   | 2272   | 13.94414 | 1.00E-05 | 0.129659 | 789  |
| HRH3   | 3537   | 5.210238 | 1.00E-05 | 0.215675 | 1125 |
| HSD11B1| 2920   | 3.240975 | 0.318972 | 0.319318 | 986  |
| HSP90AA1| 1027   | 47.18547 | 0.616805 | 0.05943  | 341  |
| HTR2A  | 3084   | 5.666719 | 1       | 0.091142 | 1050 |
| HTR2C  | 2100   | 4.290852 | 1.00E-05 | 0.093425 | 651  |
| HTR6   | 1767   | 5.128073 | 1       | 0.093766 | 589  |
| HTR7   | 1119   | 8.742291 | 0.4629  | 0.099466 | 406  |
| IGF1R  | 1538   | 5.617079 | 1.00E-05 | 0.140429 | 476  |
| INSR   | 1034   | 8.139475 | 0.754774 | 0.001    | 347  |
| ITK    | 1029   | 9.658952 | 1       | 0.001    | 352  |
| JAK2   | 1604   | 6.02122  | 0.612256 | 0.001    | 542  |
| JAK3   | 1247   | 10.2276  | 1.00E-05 | 0.001    | 418  |
| KCNH2  | 3960   | 8.742291 | 0.4629  | 0.099466 | 1315 |
| KDR    | 4303   | 5.306979 | 0.883249 | 0.001    | 1446 |
| KIT    | 1200   | 7.080043 | 1.00E-05 | 0.001    | 422  |
| LCK    | 1980   | 5.69277  | 1.00E-05 | 0.108415 | 672  |
| MAOB   | 1256   | 5.428135 | 1       | 0.140272 | 432  |
| MAPK14 | 2941   | 5.884073 | 1       | 0.066411 | 1006 |
| MAPK8  | 1399   | 7.783612 | 1.00E-05 | 0.001    | 449  |
| MAPK9  | 1141   | 24.04853 | 1       | 0.040438 | 362  |
| MAPKAPK2| 1201   | 6.007804 | 0.615441 | 0.001    | 419  |
| MC4R   | 1860   | 4.698876 | 1       | 0.001    | 656  |
| Gene   | Chr  | Log2FoldChange | P.Value | Adj.P.Value | Log2FC | FC  |
|--------|------|----------------|---------|-------------|--------|-----|
| MCHR1  | 2375 | 8.742291       | 0.4629  | 0.099466    | 823    |
| MET    | 2113 | 4.771763       | 1       | 0.001       | 698    |
| MMP1   | 1549 | 4.923501       | 1.00E-05| 0.001       | 522    |
| MMP13  | 1401 | 3.166429       | 1.00E-05| 0.001       | 455    |
| MMP2   | 1782 | 5.953406       | 1       | 0.05575     | 655    |
| MMP3   | 1214 | 5.432160       | 0.728287| 0.08936     | 352    |
| MMP9   | 1474 | 5.002715       | 1       | 0.107064    | 525    |
| MTR    | 2204 | 5.068441       | 1       | 0.149466    | 703    |
| NPY5R  | 1120 | 5.032776       | 1       | 0.001       | 396    |
| NR3C1  | 1882 | 4.761794       | 1       | 0.108338    | 640    |
| NTRK1  | 965  | 14.69707       | 1.00E-05| 0.034661    | 331    |
| OPRD1  | 3966 | 4.852147       | 0.719189| 0.001       | 1371   |
| OPR1   | 2717 | 4.902168       | 1.00E-05| 0.50189     | 950    |
| OPRL1  | 1020 | 4.990308       | 0.420702| 0.001       | 324    |
| OPRM1  | 4391 | 4.903534       | 0.772196| 0.164865    | 1439   |
| P2RX7  | 1562 | 5.513403       | 0.834737| 0.089016    | 531    |
| PARRP1 | 1166 | 5.968964       | 1       | 0.17968     | 417    |
| PDE5A  | 1094 | 5.22403       | 0.874205| 0.001       | 373    |
| PDGFRB | 1143 | 7.055928       | 1       | 0.108338    | 366    |
| PGR    | 1186 | 9.204831       | 1.00E-05| 0.001       | 450    |
| PIK3CA | 1700 | 5.401167       | 1       | 0.096607    | 608    |
| PIM1   | 1599 | 12.66678       | 1       | 0.051182    | 527    |
| PIM2   | 985  | 10.88732       | 1       | 0.001       | 351    |
| PLK1   | 2831 | 1.801875       | 0.347373| 0.049878    | 1006   |
| PPARA  | 1551 | 11.81648       | 1       | 0.139596    | 535    |
| PPARD  | 990  | 4.803409       | 0.882041| 0.246855    | 305    |
| PPARG  | 1991 | 18.15781       | 1       | 0.180451    | 711    |
| PRKACA | 963  | 3.267429       | 1.00E-05| 0.097267    | 308    |
| PRKCD  | 1127 | 12.64083       | 1       | 0.048781    | 398    |
| PTGDR2 | 1032 | 6.201784       | 1       | 0.001       | 343    |
| PTGS2  | 2120 | 5.822044       | 1.00E-05| 0.140101    | 741    |
| PTPN1  | 1106 | 10.71272       | 0.517977| 0.023416    | 347    |
| REN    | 1848 | 4.820252       | 0.948119| 0.20463     | 650    |
| ROCK1  | 1213 | 6.778567       | 1       | 0.108518    | 416    |
| ROCK2  | 1337 | 4.585667       | 0.255628| 0.001       | 423    |
| S1PR1  | 1027 | 5.609678       | 1.00E-05| 0.001       | 334    |
| SCN9A  | 1243 | 5.04573       | 1.00E-05| 0.150391    | 411    |
| SIGMAR1| 2201 | 5.78814        | 1       | 0.089305    | 685    |
| SLC6A2 | 2891 | 5.231507       | 1.00E-05| 0.204984    | 974    |
| SLC6A3 | 3778 | 6.533793       | 1       | 0.173887    | 1228   |
| SRC    | 2295 | 3.912235       | 0.358814| 0.096521    | 791    |
| TACR1  | 1816 | 5.701923       | 1.00E-05| 0.122626    | 596    |
| TRPV1  | 2238 | 4.973172       | 1.00E-05| 0.001       | 742    |
| VDR    | 1931 | 4.74124        | 0.794567| 0.186727    | 654    |