|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Iteration | SNP count | Description | 1- Mean reliability | R2 of Prediction ~ COI | R2 of Prediction ~ COI for ‘training’ individuals only | PredictionvGT.jpg | PredictionvPEV.jpg |
| 1 | 95 | Summary SNPs from iteration 4 | 0.9658684 | 0.3363477 |  |  |  |
| 2 | 85 | Used Iter\_4 gwas\_2 Top summary 85 sigSNPs (all normal filters + MQ=15) | 0.9494746 | 0.2732518 |  |  |  |
| 3 | 211 | Used Iter\_4 gwas Top P-val<0.0001 211 sigSNPs (all normal filters + MQ=15 + HWE1-e4) | 1.1442535 | 0.2954709 |  |  | 5 |
| 4 | 238 | Used Iter\_4 gwas\_2 Top P-val<0.0001 238 sigSNPs (all normal filters + MQ=15) (note, only 7 SNPs lesser than the gwas run with no MQ=15 filter) | 1.1013940 | 0.3451213 | 0.3556766 |  |  |
| 5 | 1000 | Used Iter\_4 gwas\_2 Top 1000 sigSNPs (all normal filters + MQ=15 + HWE 1e-4) | 0.9763308 | 0.7706185 | 0.4532515 |  |  |
| 6 | 70 | Summary SNPs from iteration 4's qualtable results, problematic SNPs removed subjectively. Results of qual table were generated from vcf of just MQ = 15 filter but sigSNPs of: Significant sites were gathered from iter\_2 gwas\_2 and their positions were placed into /data/meta (Iteration 1 above) | 1.0137898 | 0.3060465 |  |  |  |
| 7 | 235 | GAPIT\_7 - Used Iter 2 gwas\_2 summary 95 SNPs (all normal filters) then removed the same 25 problematic SNPs based on depth, HWE, blast res, MAF and replaced with the genotype as phenotype gwas | 0.815887 | 0.3776005 |  |  |  |
| 8 | 11290 | Thinned iter\_4/data/catvcf/Mq\_filt\_cat.vcf.gz" --maf 0.1 --thin 10000 --recode --recode-INFO-all --stdout (MQ=15 catvcf). Used that for kinship matrix.    Concatenated without -a | 0.8170614 | 0.7239177 |  |  |  |
| 9 | 200 | Reran GAPIT\_7 but with HWE 1e-4  Kept 200 out of a possible 235 Sites | 0.8884811 | 0.4084542 |  |  |  |
| 10 | 11504 | Thinned iter\_4/data/catvcf/Mq\_filt\_cat.vcf.gz" --maf 0.1 --thin 10000 --recode --recode-INFO-all --stdout (MQ=15 catvcf). Used that for kinship matrix.    Concatenated with -a and -D | 0.513678 | 0.8793138 | 0.220512 |  |  |
| 11 | 6494 | CHR08 19 million -> 19.5 million with a MAF of 0.05 and use all of it for genomic prediction + GAPIT\_4 sigSNPs | 0.8489891 | 0.529368 |  |  |  |
| 12 | 7943 | Thinning the dataset - 1 SNP in 3000BP window -> using that in GAPIT + GAPIT\_4 sigSNPs | 0.5971376 | 0.8072236 | 0.25801528 |  |  |
| 13 | 4210 | Thinning the dataset - 1 SNP in 6000BP window -> using that in GAPIT | 0.636554 | 0.7059185 | 0.29105787 |  |  |
| 14 | 129366 | Blasted an old Dartseq onto reference genome, extracted positions from imputed beagle vcf. | 0.5038456 | 0.8469278 |  |  |  |
| 15 | 7942 | Grabbed the P<0.0001 SNPs of COI from a parent gwas (GWAS\_5, ran with breeding value of COI and RGR). (209 sigSNPs) Used previous thinned dataset for kinship, note not filtered for HWE 1e-5 but MQ=15 along with others | 0.5969571 | 0.8378245 |  |  |  |
| 16 | 14916 | Thinning the dataset - 1 SNP in 15000BP window -> using that in GAPIT (14707 thinned) | 0.5300132 | 0.8822740 | 0.30724723 |  |  |
| 17 | 2377 | Top P<0.001 SNPs from iter\_4/gwas\_2 | 0.6279925 | 0.8897313  For only cross val individuals: 0.5069734 |  |  |  |
| 18 | 9981 | Rerun DArT gBLUP with --thin 100 | 0.7344284 | 0.6243695 |  |  |  |
| 19 | 675 | Top 2000 marker SNPs of parents from breeding value gwas\_5 | 0.7895513 | 0.3802566 |  |  |  |
| 20 | 524 | Run HWE 1e-5 GWAS of parents BV -> sigSNPs -> seedling positions extracted with another HWE 1e-5 filter - Top 2000 marker SNPs of parents | 0.7590057 | 0.3466395 |  |  |  |
| 21 | 9657 | Run HWE 1e-5 GWAS of parents BV -> sigSNPs -> seedling positions extracted with another HWE 1e-5 filter - Top P<0.001 SNPs of parents (1933 SNPs)    Plus thinned vcf from GAPIT\_12 | 0.6850976 | 0.7190180 |  |  |  |
| 22 | 2377 | GAPIT\_17  Increasing testing vs training gBLUP phenotypes  1. 100:420  2. 200:320  3. 300:220 | * 1. 0.6370096   2. 0.7374813   3. 0.7196880 | * 1. 0.8598871   2. 0.8213648   3. 0.7398323 |  | E.g., of 3 | E.g., of 3 |
| 23 | 2377 | Generate genomic predictions for COI for the Qld samples, and the small batch of broad leaved melaleucas (2377 SNPs). Used SNPs and genotype of GAPIT 17.    All genotypes of seedlings were used as training, outgroups are testing | 0.6205789 |  |  |  |  |
| 24 | 7733 | Gapit\_12 of only thinned SNPs | 0.9025686 | 0.6313017 | -0.01861275 |  |  |
| 25 | 2377 | Predicting Gapit\_17 onto outlier populations (152 testing, 368 training) | 0.627657 | 0.826054 |  |  |  |
| 26 | 2377 | Using GAPIT\_17 sigSNPs (seedlings (2377 SNPs) on parents (HWE 1e-4, MQ 15) to predict BV (1926 SNPs). Cross-validating these FID BV with non-genotyped seedling BV |  | 0.245 (comparing BV and predicted) |  |  |  |
| 27 | 1377 | Using GAPIT\_17 sigSNPs thinned by 100bp (originally 2377 SNPs) | 0.4979962 | Original: 0.8740454  For only cross val individuals: 0.5121157 |  |  |  |
| 28 | 869 | Using GAPIT\_17 sigSNPs thinned by 500bp (originally 2377 SNPs) | 0.3888259 | 0.8687896  For only cross val individuals: 0.5206568 |  |  |  |
| 29 | 689 | Using GAPIT\_17 sigSNPs thinned by 1000bp (originally 2377 SNPs) | 0.3655729 | 0.8560664 | 0.5421888 |  |  |
|  |  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | NA, p666c7140, p1bff81d0 | 471 | 29 | 0.343675 | 0.833808 | 0.578471 | | 2 | p0c317820, p666c7140, pd135bc30 | 449 | 29 | 0.375462 | 0.834274 | 0.710689 | | 3 | pe59c2ef0, p60668f30, pa9710de0 | 432 | 29 | 0.364977 | 0.841051 | 0.576463 | | 4 | p831064a0, NA, p9c3c9590 | 460 | 29 | 0.369848 | 0.848949 | 0.608306 | | 5 | p9c3c9590, p831064a0, p666c7140 | 444 | 29 | 0.371477 | 0.841169 | 0.665881 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 29 | 0.365 | 0.0125 | 0.84 | 0.0062 | 0.628 | 0.0587 | | | | | | |
|  |  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | nrow(my\_Y\_train) | iteration | V2 | R2 | R2\_testing | | 1 | 407 | 29 | 0.392413 | 0.825987 | 0.682376 | | 2 | 413 | 29 | 0.379421 | 0.826234 | 0.615717 | | 3 | 421 | 29 | 0.367029 | 0.828752 | 0.692911 | | 4 | 426 | 29 | 0.377945 | 0.835843 | 0.618009 | | 5 | 421 | 29 | 0.351635 | 0.841375 | 0.595879 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 29 | 0.374 | 0.0153 | 0.832 | 0.00674 | 0.641 | 0.0436 | | | | | | |
| 30 | 559 | Using GAPIT\_17 sigSNPs thinned by 2000bp (originally 2377 SNPs) | 0.3366699 | 0.8471408 | 0.5191254 |  |  |
| 31 | 1890 | Using GAPIT\_17 sigSNPs with HWE filter of 1e-4 | 0.6191737 | 0.8005745 | 0.4948388 |  |  |
| 32 | 1003 | Using GAPIT\_17 sigSNPs with HWE filter of 1e-4, thin 200 | 0.499978 | 0.803286 | 0.5122733 |  |  |
| 33 | 15566 | Merging iteration 16 (background SNPs) & 29 (sigSNPs) | 0.3353217 | 0.9408497 | 0.45035318 |  |  |
| 34 | 16587 | Merging iteration 16 (background SNPs) & 31 (sigSNPs) | 0.3420504 | 0.9165391 | 0.44450131 |  |  |
| 35 (NLR/exact/thinned) | 12151 | # Annotated NLR sourced from Jason  # Thinned by 100 and removed overlapping positions | 0.853637758636319 | 0.5781 | -0.018645 |  |  |
| 36 (NLR/exact/unthinned) | 254184 | # Annotated NLR sourced from Jason | 0.843960667473025 | 0.2465 | -0.01936 |  |  |
| 36 | 689 | # Rerunning GAPIT\_29 but replacing the top 20 SNPs with SNPs unique to the relevant sequence | 0.3663517 | 0.8565975 | 0.537653 |  |  |
| 37 | 530673 | PPR locations of Chr3 Eucalyptus obliqua | 0.8184123 | 0.8169468 | -0.01052233 |  |  |
| 38 | 128466 | GAPIT\_37 with dataset filtered for trait Res\_Sus and P-val <= 0.05 | 0.8164318 | 0.7548855 | -0.01265301 |  |  |
| 39 | 2552 | Genotyping model (DArTag) using all 3000 randomly selected markers from the provided 9000 markers | 0.3291888 | 0.9250719 | 0.5756896 |  |  |
| 40 | 996 | As above, but with only those selected for MR markers | 0.6298315 (50 training)  0.5913035 (150 training) | 0.8102004 (50 training)  0.7884592 (150 training) | 0.5697814 (50 training)  0.6288304 (150 training) | (150 individuals) |  |
|  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | p666c7140, pbb515550, p831064a0 | 447 | 40 | 0.592163 | 0.807241 | 0.757102 | | 2 | p93a6ddc0, p72bbeb40, pbb515550 | 428 | 40 | 0.573341 | 0.808685 | 0.692486 | | 3 | p0c317820, pd135bc30, pb659e150 | 426 | 40 | 0.605301 | 0.792797 | 0.672467 | | 4 | p831064a0, p8a9b6b20, pd135bc30 | 437 | 40 | 0.479109 | 0.807013 | 0.688475 | | 5 | p9c3c9590, pa9710de0, p831064a0 | 434 | 40 | 0.442405 | 0.797348 | 0.598634 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 40 | 0.538 | 0.073 | 0.803 | 0.0071 | 0.682 | 0.0566 | | | | | | | |
|  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | nrow(my\_Y\_train) | iteration | V2 | R2 | R2\_testing | | 1 | 420 | 40 | 0.580983 | 0.795379 | 0.605226 | | 2 | 421 | 40 | 0.595228 | 0.802555 | 0.670617 | | 3 | 426 | 40 | 0.54843 | 0.791277 | 0.613718 | | 4 | 411 | 40 | 0.56675 | 0.791528 | 0.709535 | | 5 | 418 | 40 | 0.560271 | 0.792871 | 0.65915 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 40 | 0.570 | 0.0182 | 0.795 | 0.00467 | 0.652 | 0.0429 | | | | | | | |
| 41 | 555 | Prediction of 555 markers from GAPIT\_39. Used to test against log10 precipitation from worldclim 0.5 degree. | 0.001050512 | 0.8587418 | 0.886103 |  |  |
| 42 | 2014 | Top 2000 MR related epistatic hits | 0.6444208 | 0.7700519 | 0.5003549 |  |  |
| 43 | 997 (50 SNPs epi, 947 MR) | Same as Run 40 but:  Bottom 50 MR SNPs replaced with epistatic non-MR SNPs | 0.5852423 (50 training)  0.7666183 (150 training) | 0.8083174 (50 training)  0.7851399 (150 training) | 0.5808438 (50 training)  0.6236592 (150 training) |  |  |
|  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | p9c3c9590, p831064a0, popnInfo-m1lg07j | 435 | 43 | 0.5867058 | 0.794144 | 0.6969092 | | 2 | p72bbeb40, p9c3c9590, p60668f30 | 426 | 43 | 0.617132 | 0.7943757 | 0.6138224 | | 3 | popnInfo-m1lg07j, p831064a0, p328a8440 | 427 | 43 | 0.6248115 | 0.8026073 | 0.7619148 | | 4 | p60668f30, p72bbeb40, pb9a5aec0 | 415 | 43 | 0.581824 | 0.7885335 | 0.6224832 | | 5 | pb9a5aec0, p8a9b6b20, popnInfo-w4ba03a | 450 | 43 | 0.580161 | 0.801403 | 0.604002 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 43 | 0.598 | 0.0212 | 0.796 | 0.0058 | 0.66 | 0.0679 | | | | | | | |
|  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Training number | iteration | V2 | R2 | R2\_testing | | 1 | 410 | 43 | 0.5506998 | 0.79217 | 0.665984 | | 2 | 415 | 43 | 0.551235 | 0.799176 | 0.681503 | | 3 | 406 | 43 | 0.5861344 | 0.782188 | 0.597503 | | 4 | 408 | 43 | 0.5803386 | 0.792993 | 0.610845 | | 5 | 414 | 43 | 0.5249972 | 0.781168 | 0.635152 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 43 | 0.559 | 0.0249 | 0.79 | 0.00768 | 0.638 | 0.0356 | | | | | | | |
| 44 | 1000 | BSLMM top 1000 SNPs | 0.6497708 | 0.7791062 | 0.3180254 |  |  |
| 45 | 500 | BSLMM top 500 SNPs | 0.3650349 | 0.8429574 | 0.5333694 |  |  |
| 46 | 997 (50 SNPs BSLMM, 947 MR) | Same as Run 40 but:  Bottom 50 MR SNPs replaced with BSLMM non-MR SNPs | 0.5446107 | 0.8148931 | 0.5665020 |  |  |
|  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | p72bbeb40, pb9a5aec0, popnInfo-w4ba03a | 434 | 46 | 0.717631 | 0.81008 | 0.659222 | | 2 | pe59c2ef0, pa9710de0, p93a6ddc0 | 431 | 46 | 0.606267 | 0.822289 | 0.57434 | | 3 | pd135bc30, popnInfo-m1lg07j, pa9710de0 | 444 | 46 | 0.591926 | 0.818081 | 0.777836 | | 4 | p1bff81d0, NA, pd135bc30 | 466 | 46 | 0.430865 | 0.813128 | 0.609337 | | 5 | p666c7140, pb659e150, p72bbeb40 | 421 | 46 | 0.555566 | 0.804349 | 0.701776 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 46 | 0.58 | 0.103 | 0.814 | 0.00696 | 0.665 | 0.0797 | | | | | | | |
|  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Training number | iteration | V2 | R2 | R2\_testing | | 1 | 415 | 46 | 0.595534 | 0.804828 | 0.648976 | | 2 | 403 | 46 | 0.544182 | 0.79565 | 0.649244 | | 3 | 413 | 46 | 0.555019 | 0.795494 | 0.676095 | | 4 | 397 | 46 | 0.580508 | 0.796251 | 0.642976 | | 5 | 409 | 46 | 0.569904 | 0.799877 | 0.603301 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 46 | 0.569 | 0.0203 | 0.798 | 0.004 | 0.644 | 0.0262 | | | | | | | |
| 47 | 997 (50 SNPs BSLMM, 50 SNPs epi, 897 MR) | Same as Run 40 but:  Bottom 100 MR SNPs replaced with 50 epistatic non-MR SNPs AND 50 BSLMM non-MR SNPs | 0.5945163 | 0.8038652 | 0.5852823 |  |  |
|  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | popnInfo-m1lg07j, pe59c2ef0, p60668f30 | 433 | 47 | 0.585091 | 0.796 | 0.554336 | | 2 | pd135bc30, popnInfo-w4ba03a, p9c3c9590 | 464 | 47 | 3.874431 | 0.802085 | 0.646146 | | 3 | p8a9b6b20, p1bff81d0, p60668f30 | 435 | 47 | 0.554299 | 0.786114 | 0.5198244 | | 4 | p9c3c9590, p328a8440, p8a9b6b20 | 440 | 47 | 0.596311 | 0.79609 | 0.603869 | | 5 | pa9710de0, p9c3c9590, p0c317820 | 438 | 47 | 0.59553 | 0.789412 | 0.613198 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 47 | 1.24 | 1.47 | 0.794 | 0.00626 | 0.587 | 0.0501 | | | | | | | |
|  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Training number | iteration | V2 | R2 | R2\_testing | | 1 | 415 | 47 | 0.627067 | 0.796 | 0.68483 | | 2 | 409 | 47 | 0.603117 | 0.774485 | 0.602206 | | 3 | 393 | 47 | 0.659114 | 0.798042 | 0.707504 | | 4 | 412 | 47 | 0.591633 | 0.791105 | 0.644536 | | 5 | 421 | 47 | 0.629625 | 0.786969 | 0.630671 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 47 | 0.622 | 0.0262 | 0.789 | 0.00935 | 0.654 | 0.0422 | | | | | | | |
| 48 | 3005 | Final 3,300 SNPs for DArTag panel; missing SNPs are climate SNPs + 1 MR SNP 15434364 | Actual reliability: 0.728 | 0.895 | 0.619 |  |  |
|  |  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Selected population | Training number | iteration | Reliability | R2 | R2\_testing | | 1 | p666c7140, p72bbeb40, | p8a9b6b20 | 436 | 48 | 0.734186 | 0.922308 | | 2 | p666c7140, pbb515550, | p93a6ddc0 | 452 | 48 | 0.737533 | 0.926412 | | 3 | popnInfo-m1lg07j, pb659e150, | pe59c2ef0 | 422 | 48 | 0.735875 | 0.907143 | | 4 | p93a6ddc0, p831064a0, | pbb515550 | 434 | 48 | 0.726917 | 0.921196 | | 5 | p93a6ddc0, popnInfo-m1lg07j, | p328a8440 | 432 | 48 | 0.754558 | 0.919852 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 48 | 0.738 | 0.0102 | 0.919 | 0.00727 | 0.662 | 0.035 | |  |  |  |  |  |
|  |  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Training number | iteration | V2 | R2 | R2\_testing | | 1 | 418 | 48 | 0.720187 | 0.898467 | 0.676102 | | 2 | 403 | 48 | 0.742457 | 0.887689 | 0.574273 | | 3 | 415 | 48 | 0.729701 | 0.899251 | 0.607729 | | 4 | 402 | 48 | 0.735619 | 0.877743 | 0.558548 | | 5 | 416 | 48 | 0.71026 | 0.909725 | 0.678139 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 48 | 0.728 | 0.0127 | 0.895 | 0.0122 | 0.619 | 0.056 | |  |  |  |  |  |
| 48.5 | 1049 | Final MR and epi SNPs for DArTag panel, 1 SNP missing as above 15434364 | Actual reliability: 0.683 | 0.875 | 0.683 |  |  |
|  |  | **Randomised Pop Run**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | 1 | p328a8440, | p9c3c9590, | p1bff81d0 | 435 | 48 | 0.692307 | | 2 | p72bbeb40, | popnInfo-w4ba03a, | popnInfo-m1lg07j | 442 | 48 | 0.687889 | | 3 | pbb515550, | p9c3c9590, | pa9710de0 | 449 | 48 | 0.679417 | | 4 | popnInfo-m1lg07j, | p72bbeb40, | p328a8440 | 421 | 48 | 0.684642 | | 5 | popnInfo-w4ba03a, | pd135bc30, | p831064a0 | 452 | 48 | 0.668776 | | 1 | p328a8440, | p9c3c9590, | p1bff81d0 | 435 | 48 | 0.692307 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 48 | 0.683 | 0.00905 | 0.875 | 0.0141 | 0.683 | 0.0516 | |  |  |  |  |  |
|  |  | **Randomised FID Run**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Training number | iteration | V2 | R2 | R2\_testing | | 1 | 417 | 48 | 0.682905 | 0.844695 | 0.632816 | | 2 | 422 | 48 | 0.682035 | 0.864739 | 0.621449 | | 3 | 423 | 48 | 0.668357 | 0.86253 | 0.721841 | | 4 | 399 | 48 | 0.685637 | 0.86568 | 0.679173 | | 5 | 412 | 48 | 0.675906 | 0.850687 | 0.663161 |   Mean of above   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | iteration | rel\_mean | rel\_sd | R2\_mean | R2\_sd | R2\_test\_mean | R2\_test\_sd | | 48 | 0.679 | 0.00692 | 0.858 | 0.00942 | 0.664 | 0.0399 | |  |  |  |  |  |
| 50 | 1000 | 1,000 SNPs with top associations P < 9.45E-06 (SNPs 316 – 1316) | Actual reliability: 0.29045 | 0.741825 | 0.488182 |  |  |
| 51 | 190473 | What if we didn't know which SNPs were significant and just annotated the genome for disease proteins? Using all NLR & NBARC genes for prediction | 0.19039 | 0.6390249 | 0.05454132 |  |  |
|  | 171318 | Same as above but NLR only | 0.1942611 | 0.6125443 | 0.05193969 |  |  |