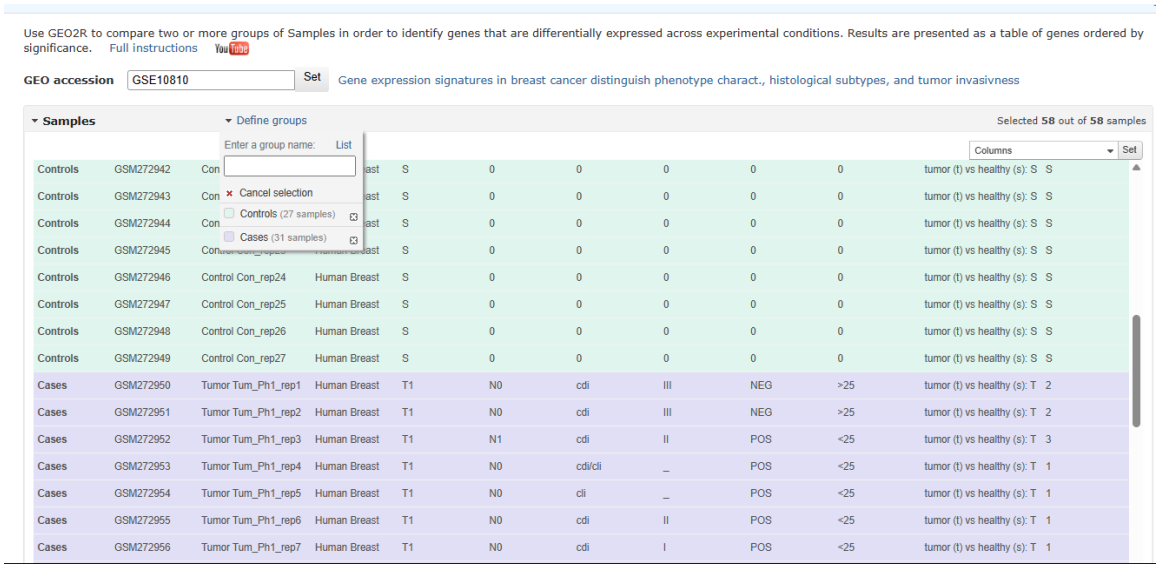
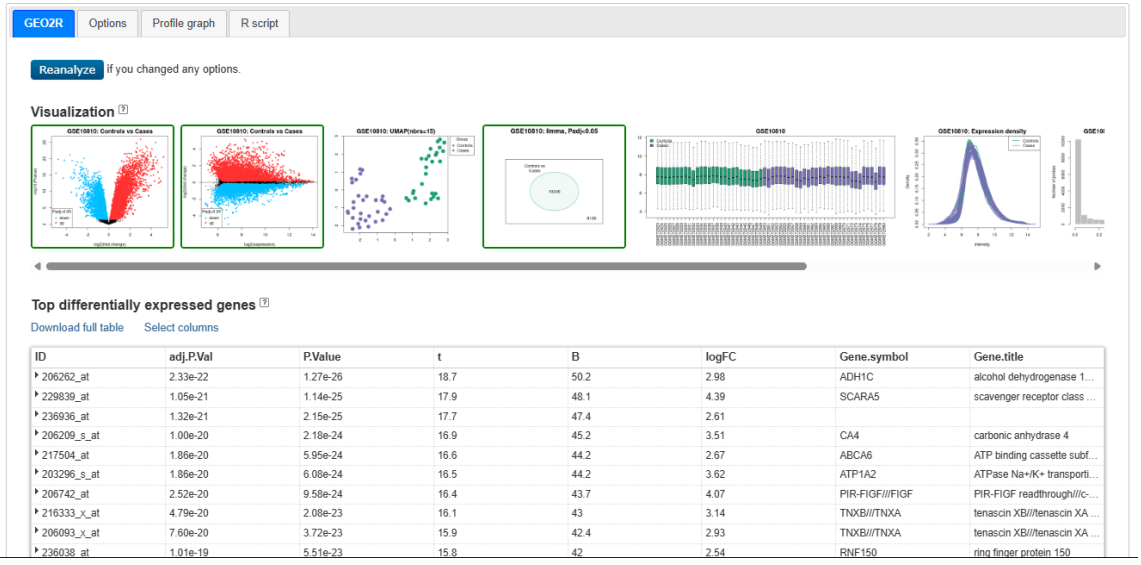
# ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ 3

## A)Ανάλυση γονιδιακής έκφρασης με το GEO2R και ανάλυση εμπλουτισμού με το gProfiler (GSE10810)

α) Ορίζω ομάδες Cases και Controls για τους ασθενείς και τους υγιείς αντίστοιχα και στην συνέχεια κάνω ανάλυση διαφορικής έκφρασης .

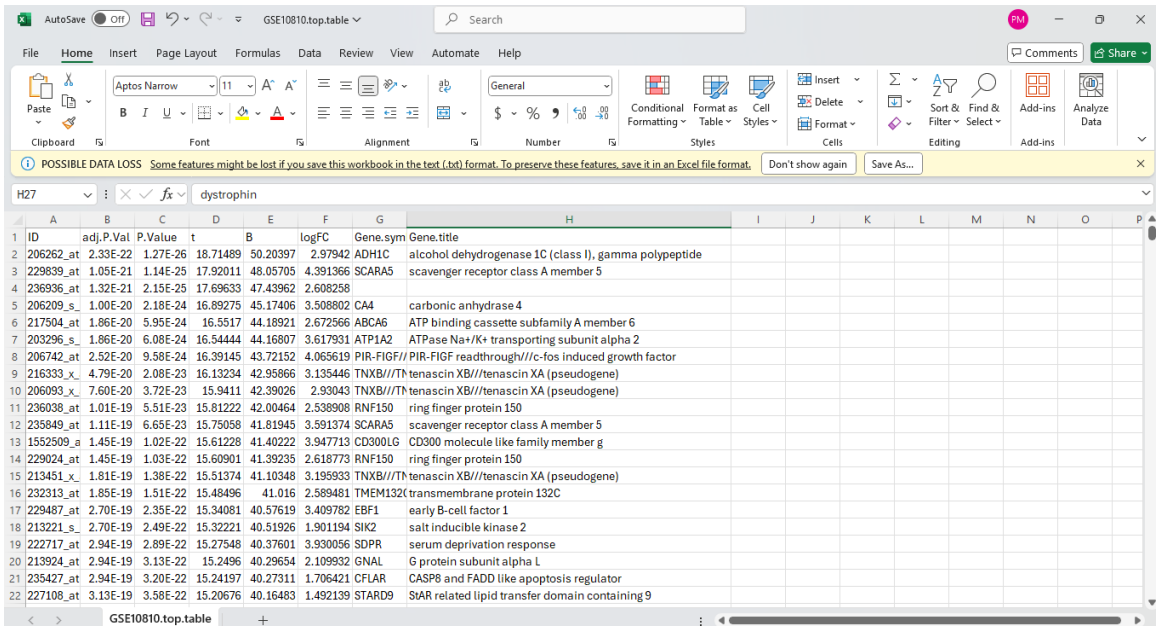



β)Αποτελέσματα της ανάλυσης:



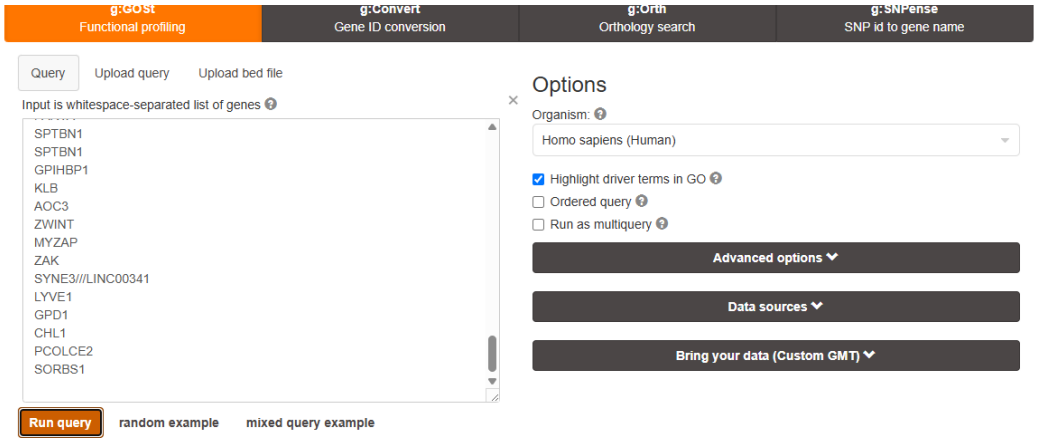

Στο παραπάνω εχει χρησιμοποιηθεί η μέθοδος διόρθωσης **Benjamini & Hochberg (False discovery rate).**

γ)Κατεβάζω τα αποτελέσματα τις ανάλυσης και επιλέγω τα top 100 διαφορικά εκφρασμένα γονίδια ($p < 0.05$ ):

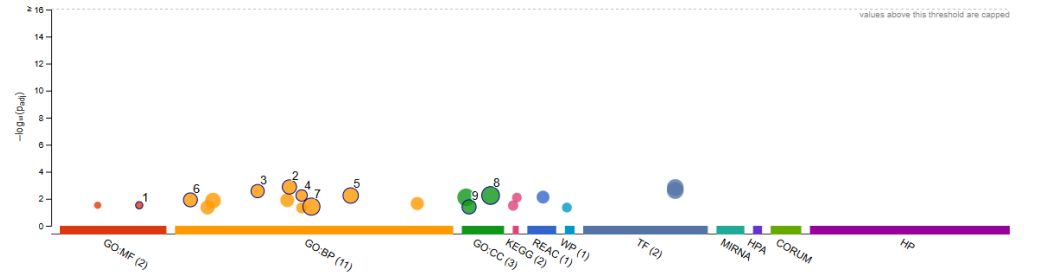


| | ID | adj.P.Val | P.Value | t | B | logFC | Gene.sym | Gene.title |
|---|---|---|---|---|---|---|---|---|
| 2 | 206262_at | 2.33E-22 | 1.27E-26 | 18.71489 | 50.20397 | 2.97942 | ADH1C | alcohol dehydrogenase 1C (class I), gamma polypeptide |
| 3 | 229839_at | 1.05E-21 | 1.14E-25 | 17.92011 | 48.05705 | 4.391366 | SCARA5 | scavenger receptor class A member 5 |
| 4 | 236936_at | 1.32E-21 | 2.15E-25 | 17.69633 | 47.43962 | 2.608258 | | |
| 5 | 206209_s_ | 1.00E-20 | 2.18E-24 | 16.89275 | 45.17406 | 3.508802 | CA4 | carbonic anhydrase 4 |
| 6 | 217504_at | 1.86E-20 | 5.95E-24 | 16.5517 | 44.18921 | 2.672566 | ABCA6 | ATP binding cassette subfamily A member 6 |
| 7 | 203296_s_ | 1.86E-20 | 6.08E-24 | 16.54444 | 44.16807 | 3.617931 | ATP1A2 | ATPase Na+/K+ transporting subunit alpha 2 |
| 8 | 206742_at | 2.52E-20 | 9.58E-24 | 16.39145 | 43.72152 | 4.065619 | PIR-FIGF// | PIR-FIGF readthrough///c-fos induced growth factor |
| 9 | 216333_x_ | 4.79E-20 | 2.08E-23 | 16.13234 | 42.95866 | 3.135446 | TNXB///TN | tenascin XB///tenascin XA (pseudogene) |
| 10 | 206093_x_ | 7.60E-20 | 3.72E-23 | 15.9411 | 42.39026 | 2.93043 | TNXB///TN | tenascin XB///tenascin XA (pseudogene) |
| 11 | 236038_at | 1.01E-19 | 5.51E-23 | 15.81222 | 42.00464 | 2.538908 | RNF150 | ring finger protein 150 |
| 12 | 235849_at | 1.11E-19 | 6.65E-23 | 15.75058 | 41.81945 | 3.591374 | SCARA5 | scavenger receptor class A member 5 |
| 13 | 1552509_a | 1.45E-19 | 1.02E-22 | 15.61228 | 41.40222 | 3.947713 | CD300LG | CD300 molecule like family member g |
| 14 | 229024_at | 1.45E-19 | 1.03E-22 | 15.60901 | 41.39235 | 2.618773 | RNF150 | ring finger protein 150 |
| 15 | 213451_x_ | 1.81E-19 | 1.38E-22 | 15.51374 | 41.10348 | 3.195933 | TNXB///TN | tenascin XB///tenascin XA (pseudogene) |
| 16 | 232313_at | 1.85E-19 | 1.51E-22 | 15.48496 | 41.016 | 2.589481 | TMEM132( | transmembrane protein 132C |
| 17 | 229487_at | 2.70E-19 | 2.35E-22 | 15.34081 | 40.57619 | 3.409782 | EBF1 | early B-cell factor 1 |
| 18 | 213221_s_ | 2.70E-19 | 2.49E-22 | 15.32221 | 40.51926 | 1.901194 | SIK2 | salt inducible kinase 2 |
| 19 | 222717_at | 2.94E-19 | 2.89E-22 | 15.27548 | 40.37601 | 3.930056 | SDPR | serum deprivation response |
| 20 | 213924_at | 2.94E-19 | 3.13E-22 | 15.2496 | 40.29654 | 2.109932 | GNAL | G protein subunit alpha L |
| 21 | 235427_at | 2.94E-19 | 3.20E-22 | 15.24197 | 40.27311 | 1.706421 | CFLAR | CASP8 and FADD like apoptosis regulator |
| 22 | 227108_at | 3.13E-19 | 3.58E-22 | 15.20676 | 40.16483 | 1.492139 | STARD9 | StAR related lipid transfer domain containing 9 |

Ύστερα πραγματοποιώ ανάλυση εμπλουτισμού στο gProfiler:



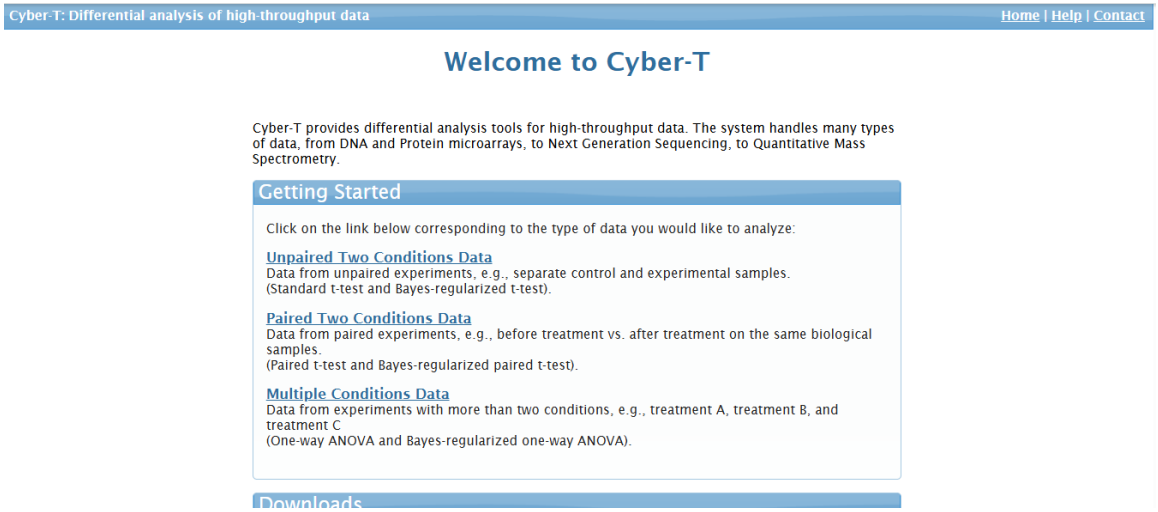

Αποτελέσματα:




| ID | Source | Term ID | Term Name | $p_{adj}$ (query_1) |
|---|---|---|---|---|
| 1 | GO:MF | GO:0052595 | aliphatic amine oxidase activity | $2.987\times10^{-2}$ |
| 2 | GO:BP | GO:0044057 | regulation of system process | $1.328\times10^{-3}$ |
| 3 | GO:BP | GO:0032868 | response to insulin | $2.688\times10^{-3}$ |
| 4 | GO:BP | GO:0046323 | D-glucose import | $5.699\times10^{-3}$ |
| 5 | GO:BP | GO:0071495 | cellular response to endogenous stimulus | $5.734\times10^{-3}$ |
| 6 | GO:BP | GO:0003012 | muscle system process | $1.199\times10^{-2}$ |
| 7 | GO:BP | GO:0048731 | system development | $3.773\times10^{-2}$ |
| 8 | GO:CC | GO:0071944 | cell periphery | $5.694\times10^{-3}$ |
| 9 | GO:CC | GO:0015629 | actin cytoskeleton | $3.981\times10^{-2}$ |

**Β) Ανάλυση γονιδιακής έκφρασης με το Cyber-T και ανάλυση εμπλουτισμού με το gProfiler**

α)Κάνω ανάλυση γονιδιακής έκφρασης με το Cyber-T




β)Υστερα επιλέγω Unpaired Two Conditions Data-> Dna Microarray Controls-> Bayes-regularized Analysis-> submit:

Results for Top 25 Genes

Help

See All Data | Download Txt

| Lab_0 | nC | nE | meanC | meanE | stdC | stdE | fold | rasdC | rasdE | bayesSDC | bayesSDE | bayesT | ba |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| hdhA | 4 | 4 | 1.297e-05 | 0.0002137 | 1.216e-05 | 2.752e-05 | 16.49 | 1.206e-05 | 4.756e-05 | 1.307e-05 | 4.344e-05 | 8.852 | 12 |
| rmf | 4 | 4 | 5.791e-05 | 0.001473 | 4.677e-05 | 0.0003353 | 25.43 | 2.781e-05 | 0.0002856 | 4.012e-05 | 0.0003326 | 8.446 | 12 |
| oppA | 4 | 4 | 0.001625 | 0.03157 | 0.000763 | 0.01033 | 19.44 | 0.000551 | 0.0007967 | 0.0007025 | 0.007331 | 8.134 | 12 |
| lysU | 4 | 4 | 0.0001808 | 0.001242 | 7.479e-05 | 0.0002785 | 6.874 | 5.39e-05 | 0.0002083 | 6.88e-05 | 0.0002602 | 7.89 | 12 |
| hdeA | 4 | 4 | 0.0002403 | 0.0008285 | 8.462e-05 | 9.901e-05 | 3.448 | 7.052e-05 | 0.0001304 | 8.304e-05 | 0.0001274 | 7.734 | 12 |
| gltD | 4 | 4 | 0.0005275 | 2.738e-05 | 0.0001284 | 1.421e-05 | -19.27 | 0.0001152 | 1.965e-05 | 0.0001307 | 1.893e-05 | -7.573 | 12 |
| oppB | 4 | 4 | 7.512e-05 | 0.001136 | 2.123e-05 | 0.0003793 | 15.12 | 3.372e-05 | 0.0001811 | 3.136e-05 | 0.0003062 | 6.894 | 12 |
| artP | 4 | 4 | 6.732e-05 | 0.000423 | 1.239e-05 | 0.0001164 | 6.283 | 3.121e-05 | 7.77e-05 | 2.694e-05 | 0.0001039 | 6.626 | 12 |
| oppD | 4 | 4 | 8.967e-05 | 0.0006547 | 2.761e-05 | 0.0002045 | 7.301 | 3.709e-05 | 0.0001017 | 3.603e-05 | 0.0001668 | 6.624 | 12 |
| pheA | 4 | 4 | 9.105e-05 | 0.0003412 | 3.778e-05 | 4.174e-05 | 3.748 | 3.796e-05 | 7.394e-05 | 4.092e-05 | 6.72e-05 | 6.36 | 12 |
| livG | 4 | 4 | 0.0001792 | 1.811e-05 | 3.318e-05 | 1.936e-05 | -9.896 | 5.243e-05 | 1.433e-05 | 4.882e-05 | 1.801e-05 | -6.193 | 12 |
| ydaR | 4 | 4 | 5.145e-05 | 0.0002615 | 2.613e-05 | 5.608e-05 | 5.083 | 2.564e-05 | 5.808e-05 | 2.793e-05 | 6.182e-05 | 6.193 | 12 |
| gdhA | 4 | 4 | 9.163e-05 | 0.0002733 | 1.519e-05 | 2.162e-05 | 2.982 | 3.768e-05 | 5.77e-05 | 3.259e-05 | 4.953e-05 | 6.127 | 12 |
| oppC | 4 | 4 | 0.0002013 | 0.001083 | 2.337e-05 | 0.0003605 | 5.38 | 5.35e-05 | 0.0001762 | 4.67e-05 | 0.0002927 | 5.947 | 12 |
| trxA | 4 | 4 | 9.053e-05 | 0.0002837 | 2.993e-05 | 4.289e-05 | 3.134 | 3.728e-05 | 5.655e-05 | 3.707e-05 | 5.525e-05 | 5.808 | 12 |
| ilvG_1 | 4 | 4 | 0.0004208 | 0.000915 | 7.547e-05 | 6.846e-05 | 2.175 | 0.0001055 | 0.0001586 | 0.0001014 | 0.0001382 | 5.767 | 12 |
| grxB | 4 | 4 | 5.952e-05 | 0.000338 | 1.924e-05 | 0.0001068 | 5.678 | 2.741e-05 | 7.085e-05 | 2.619e-05 | 9.514e-05 | 5.644 | 12 |
| rpoE | 4 | 4 | 0.0001705 | 0.0004348 | 4.796e-05 | 7.288e-05 | 2.55 | 4.642e-05 | 8.487e-05 | 5.086e-05 | 8.636e-05 | 5.273 | 12 |

**Γ) GWAS ανάλυση και μετα-ανάλυση με το PLINK και ανάλυση εμπλουτισμού με το gProfiler**

- Άφου κατεβάσω το PLINK v.1.9 και τα σύνολα δεδομένων που δίνονται απο την εκφώνηση εκτελώ τις παρακάτω εντολές:

```
PS C:\Users\petri\Desktop\bioplhroforiki_II-main> wsl
abstract@MichaelPetrisis:/mnt/c/Users/petri/Desktop/bioplhroforiki_II-main$ ./plink --file toy --make-bed --out toy_bin
PLINK v1.9.0-b.7.8 64-bit (15 Jun 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to toy_bin.log.
Options in effect:
  --file toy
  --make-bed
  --out toy_bin

7515 MB RAM detected; reserving 3757 MB for main workspace.
.ped scan complete (for binary autoconversion).
Performing single-pass .bed write (2 variants, 2 people).
--file: toy_bin-temporary.bed + toy_bin-temporary.bim + toy_bin-temporary.fam
written.
2 variants loaded from .bim file.
2 people (2 males, 0 females) loaded from .fam.
2 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Total genotyping rate is 0.75.
2 variants and 2 people pass filters and QC.
Among remaining phenotypes, 1 is a case and 1 is a control.
--make-bed to toy_bin.bed + toy_bin.bim + toy_bin.fam ... done.
```

Δημιουργία toy_bin.bed, toy_bin.fam και toy_bin.bim.

```
abstract@MichaelPetrisis:/mnt/c/Users/petri/Desktop/bioplhroforiki_II-main$ ./plink --bfile toy_bin  --assoc --out assoc
_results
PLINK v1.9.0-b.7.8 64-bit (15 Jun 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to assoc_results.log.
Options in effect:
  --assoc
  --bfile toy_bin
  --out assoc_results

7515 MB RAM detected; reserving 3757 MB for main workspace.
2 variants loaded from .bim file.
2 people (2 males, 0 females) loaded from .fam.
2 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Total genotyping rate is 0.75.
2 variants and 2 people pass filters and QC.
Among remaining phenotypes, 1 is a case and 1 is a control.
Writing C/C --assoc report to assoc_results.assoc ... done.
```

Δημιουργία Assoc results

- Μέτα-ανάλυση των demo_Plink 1&2&3.txt

```
abstract@MichaelPetrisis:/mnt/c/Users/petri/Desktop/bioplhroforiki_II-main$ ./plink --meta-analysis demo_PLINK_1.txt dem
o_PLINK_2.txt  demo_PLINK_3.txt +  logscale no-allele report-all
PLINK v1.9.0-b.7.8 64-bit (15 Jun 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to plink.log.
Options in effect:
  --meta-analysis demo_PLINK_1.txt demo_PLINK_2.txt demo_PLINK_3.txt + logscale no-allele report-all

7515 MB RAM detected; reserving 3757 MB for main workspace.
--meta-analysis: 5 variants processed; results written to plink.meta .
```

| CHR | BP | SNP | N | P | P(R) | OR | OR(R) | Q | I |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 169093837 | rs2954939 | 3 | 0.991 | 0.991 | 1.0003 | 1.0003 | 0.4863 | 0.00 |
| 2 | 169095689 | rs12619614 | 3 | 0.9688 | 0.9688 | 1.0005 | 1.0005 | 0.7636 | 0.00 |
| 2 | 169095851 | rs13415004 | 3 | 0.7735 | 0.7735 | 1.0045 | 1.0045 | 0.5582 | 0.00 |
| 2 | 169095873 | rs2724164 | 3 | 0.7017 | 0.7017 | 0.9909 | 0.9909 | 0.7663 | 0.00 |
| 2 | 169097055 | rs11681374 | 3 | 0.9292 | 0.9292 | 1.0027 | 1.0027 | 0.9385 | 0.00 |

- Ο τελικός φάκελος με τα όλα τα αρχεία είναι ο παρακάτω: