**ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ ΙΙ**

**ΕΡΓΑΣΤΗΡΙΑΚΕΣ ΑΣΚΗΣΕΙΣ Ε.Ε 2024-2025**

**ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ 3**

**Α) Ανάλυση γονιδιακής έκφρασης με το GEO2R και ανάλυση εμπλουτισμού με το gProfiler**

α) Μεταβείτε στην ιστοσελίδα του GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/> )

β) Πραγματοποιήστε ανάλυση διαφορικής έκφρασης (Differential Expression Analysis) για την μελέτη **GSE10810**, ορίζοντας σωστά τις ομάδες cases (ασθενείς) και controls (υγιείς) και επιλέγοντας την μέθοδο διόρθωσης **Benjamini & Hochberg (False discovery rate)**.Στην συνέχεια, κατεβάστε τα αποτελέσματα της ανάλυσης.

γ) Επιλέξτε τα top 100 διαφορικά εκφρασμένα γονίδια (*p < 0.05* ) από τα αποτελέσματα και πραγματοποιήστε ανάλυση εμπλουτισμού στο gProfiler (<https://biit.cs.ut.ee/gprofiler/gost>)

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

α) Μεταβείτε στην ιστοσελίδα του Cyber-T (<http://cybert.ics.uci.edu/>)

β) Χρησιμοποιήστε το example datasets όπως φαίνονται στην εικόνα και πραγματοποιήστε την ανάλυση με το Cyber-T

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.

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

α) Κατεβάστε το PLINK v.1.9 (Stable, Linux 64-bit) από την ακόλουθη διεύθυνση: (<https://www.cog-genomics.org/plink/1.9/>)

β) Κατεβάστε τα ακόλουθα σύνολα δεδομένων από την διεύθυνση : <https://github.com/gmanios/bioplhroforiki_II/tree/main/PLINK_DEMO_data>

γ) Εκτελέστε την ακόλουθη εντολή στο PLINK για να πραγματοποιήσετε μετα-ανάλυση με τα τις μελέτες που κατεβάσατε στο προηγούμενο βήμα:

#Preprocessing

./plink --file toy --make-bed --out toy\_bin

#Association Test

./plink --bfile toy\_bin --assoc --out assoc\_results

#Meta-analysis

./plink --meta-analysis demo\_PLINK\_1.txt demo\_PLINK\_2.txt demo\_PLINK\_3.txt + logscale no-allele report-all

**Παραδοτέα εργασία**

**Στην παραδοτέα εργασία καλείστε να επαναλάβετε τις αναλύσεις που δείξαμε στην εργαστηριακή άσκηση. Θα πρέπει στην τελική αναφορά (docx/pdf αρχείο):**

1. **Να δείξετε αναλυτικά πως φτάσατε στα αποτελέσματα των αναλύσεων (με σύντομες επεξηγήσεις, screenshots κλπ.)**
2. **Να ερμηνεύσετε σωστά τα αποτελέσματα των αναλύσεων**

**Η εργασία θα πρέπει να υποβληθεί στα προσωπικά σας GitHub repositories, όπου θα είμαι contributor (gmanios). Οποιαδήποτε εργασία σταλθεί μετά το πέρας της προθεσμίας, δεν θα γίνεται αποδεκτή.**

**Α) Ανάλυση γονιδιακής έκφρασης με το GEO2R και ανάλυση εμπλουτισμού με το gProfiler**

α) Μεταβείτε στην σελίδα της GEO (<https://www.ncbi.nlm.nih.gov/geo/>). Βρείτε μια case-control μελέτη στην οποία οι δύο ομάδες αναφέρονται ξεκάθαρα.

β) Με την μελέτη που βρήκατε στην GEO, πραγματοποιήστε ανάλυση διαφορικής έκφρασης στο GEO2R, ορίζοντας σωστά τα cases και τα controls, και κατεβάστε τα αποτελέσματα της ανάλυσης.

γ) Επιλέξτε τα top 100 διαφορικά εκφρασμένα γονίδια (*p < 0.05* ) από τα αποτελέσματα και πραγματοποιήστε ανάλυση εμπλουτισμού στο gProfiler (<https://biit.cs.ut.ee/gprofiler/gost>)

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

***ΠΡΟΣΟΣΧΗ : Για αυτή την άσκηση, θα πρέπει να εκτελέσετε το PLINK σε περιβάλλον Linux (είτε με WSL, Cygwin, VirtualBox, κλπ.)***

α) Κατεβάστε το PLINK v.1.9 (Stable, Linux 64-bit) από την ακόλουθη διεύθυνση: (<https://www.cog-genomics.org/plink/1.9/)>

β) Κατεβάστε τα ακόλουθα σύνολα δεδομένων από την διεύθυνση : <https://github.com/gmanios/bioplhroforiki_II/tree/main/PLINK_EXAMPLE_data>

γ) Πραγματοποιήστε μετα-ανάλυση με το PLINK και ανάλυση εμπλουτισμού με τα στατιστικά σημαντικά SNPs (*p < 1e-8* ) στο gProfiler. Ερμηνεύστε και σχολιάστε τα αποτελέσματα της μετα-ανάλυσης και της ανάλυσης εμπλουτισμού.