דו"ח פרויקט סופי – ניתוח סיבתיות

**1. נתונים ורקע**

בפרויקט זה השתמשנו בתוצאות של מטלה 2 (QTL analysis) ושל מטלה 3 (eQTL analysis). מטלות אלו זיהו קשרים בין לוקוסים גנומיים לתכונות כמותיות (QTLs), וכן בין אותם לוקוסים לרמות ביטוי גנים (eQTLs). כאן שילבנו את התוצאות כדי ליצור שלשות (QTL–גן–פנוטיפ), שבהן גם הגן וגם הפנוטיפ מראים קשר עם אותו לוקוס או לוקוס סמוך. מתוך התוצאות נבחרו 10 שלשות מובילות (בהתבסס על מובהקות ומרחק גנומי) לבדיקת סיבתיות.

**2. שיטות**  
**2.1 . מודלים שנבדקו**  
שלושה מודלים אפשריים נבדקו עבור כל שלשה:

* M1: QTL → Gene → Phenotype (gene mediates the effect of the QTL on the phenotype)
* M2: QTL → Phenotype → Gene (phenotype drives changes in gene expression)
* M3: QTL → Gene and QTL → Phenotype independently (pleiotropy).

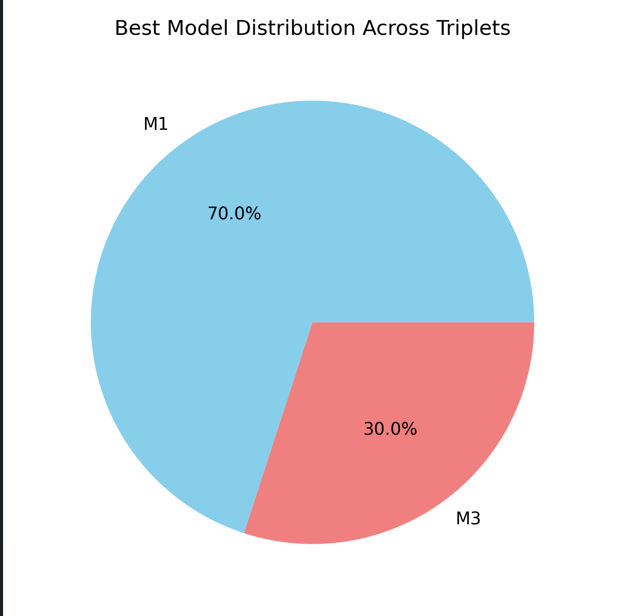
**2.2. מבחן סיבתיות**  
עבור כל שלשה חישבנו לוג-סבירות תחת שלושת המודלים באמצעות רגרסיה לינארית. המודל עם הסבירות הגבוהה ביותר נחשב להתאמה הטובה ביותר.

**2.3 . מבחן פרמוטציות**  
לצורך בדיקת מובהקות סטטיסטית בוצעו מבחני פרמוטציות. וקטור הגנוטיפ (L) עבר ערבוב בין פרטים, ובכך נשבר המבנה הסיבתי אך נשמרו ההתפלגויות. עבור כל פרמוטציה, ההפרשים בלוג-סבירות (M1–M2, M1–M3) חושבו מחדש.

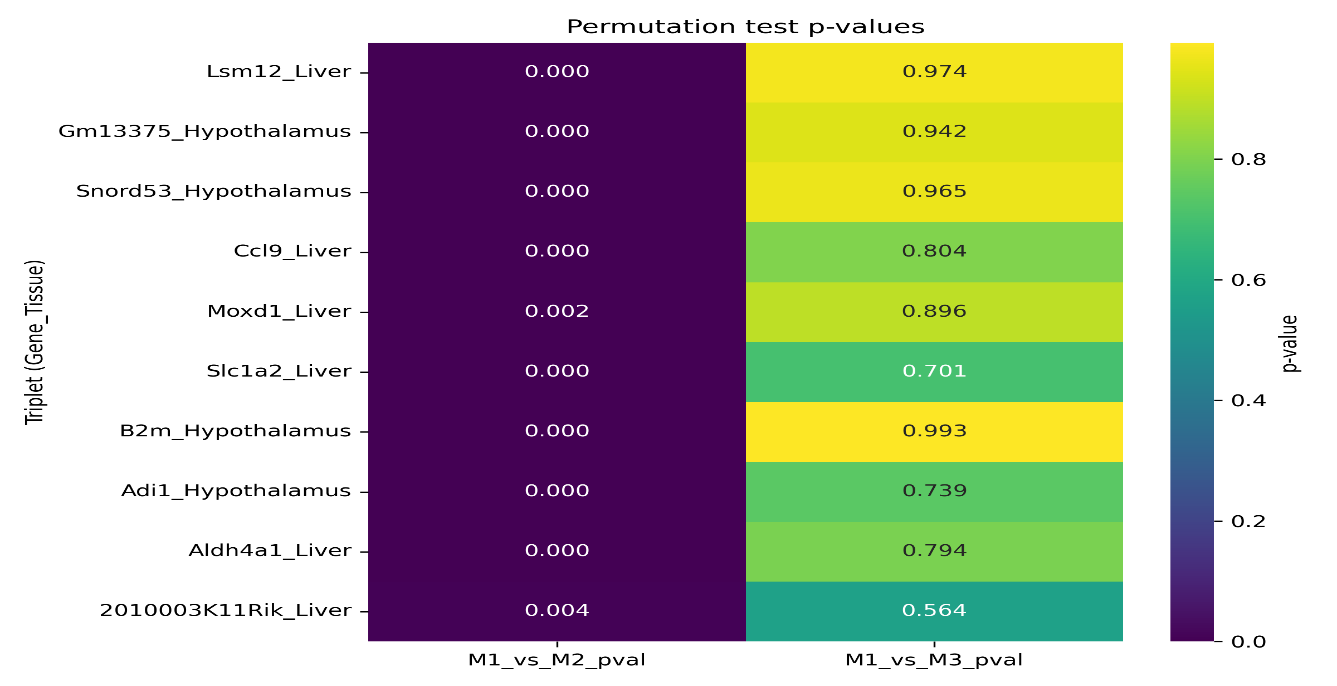
**3**. **תוצאות**  
**3.1 . טבלת סיכום**  
הסיכום המלא של כל 10 השלשות נמצא בקובץ 'causality\_results.csv'. להלן שורות מייצגות (הערכים קוצצו להמחשה):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Triplet | Gene | Tissue | SNP | Best Model | Significance (p-values) |
| 1 | Lsm12 | Liver | rs3685813 | M1 | M1 vs M2: p<0.001; M1 vs M3: ns |
| 2 | Gm13375 | Hypothalamus | rs13476379 | M1 | M1 vs M2: p<0.001; M1 vs M3: ns |

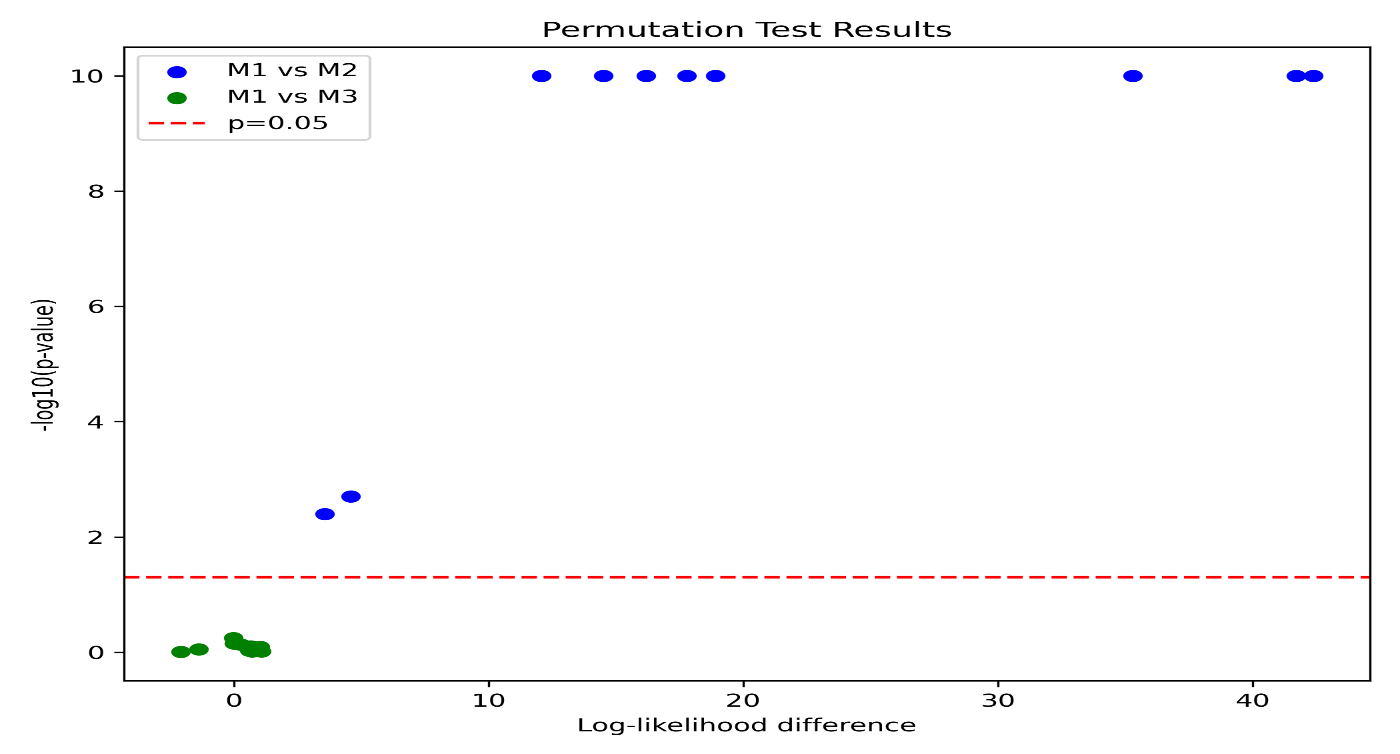
**3.2. התפלגות מודלים**  
 70% מהשלשות הוסברו בצורה הטובה ביותר על ידי מודל ,M1בעוד 30% הוסברו בצורה הטובה ביותר על ידי מודל ,M3 מודל M2 לא נבחר כמיטבי באף מקרה.



**3.3. מובהקות מבחן פרמוטציות**  
ערכי ה-p הראו תמיכה חזקה במודל M1 לעומת M2 (p < 0.01 ברוב המקרים). עבור M1 לעומת M3, ערכי ה-p היו גבוהים בדרך כלל (>0.5), מה שמרמז שאין הבדל מובהק.



**3.4. גרף וולקנו**  
גרף הוולקנו מדגיש את עוצמת ההבדלים בלוג-סבירות ואת מובהקותם. עבור M1 לעומת M2 נמצאו הבדלים גדולים ומשמעותיים, בעוד שלעומת M3 כמעט לא נמצאו הבדלים.



**4. דיון**  
מרבית השלשות תומכות במודל M1, המציין שביטוי גנים מתווך את השפעת ה-QTL על הפנוטיפים. הדבר עולה בקנה אחד עם הציפיות הביולוגיות להסדרת ביטוי גנים. במקרים מסוימים נמצא תמיכה במודל M3, המעיד על השפעות פליאוטרופיות של QTLs. מודל M2 לא נתמך, מה שמחזק את הכיוון

**5. מסקנות**  
ניתוח הסיבתיות מאשר כי התיווך דרך ביטוי גנים הוא המודל השולט בשלשות הנבחרות. מבחני הפרמוטציות אישרו את המובהקות של M1 מול M2, אך לא מול M3. המשמעות היא שפליאוטרופיה נותרה הסבר חלופי סביר במקרים מסוימים. תוצאות אלו מספקות גנים מועמדים להמשך ולידציה ביולוגית.