

## תרגיל 4 - הכנה להאקתון

גל סזנה, נעה מרגוליס, יואל מרקו, איתן סמסון ועדי יפרואימסקי

22 בינואר 2025

### חברי הקבוצה

☆ גל סזנה - 318510633

☆ נעה מרגוליס - 208635334

☆ עדי יפרואימסקי - 326045515

☆ יואל מרקו - 325245280

☆ איתן סמסון - 214935165

### רקע

מחלות מיילופרליפרטיביות (MPNs) הן סוג של סרטן דם כרוני הנגרם ממוטציות גנטיות סומטיות בתאי גזע המטופואטיים (HSPs). למרות שהמחלה נחקרה רבות בעבר, עדיין חסר מידע אודות המוטציות המובילות הגורמות להתפתחות המחלה, והתזמון שבו הן מופיעות לאורך חייו של החולה. כחלק מתהליך המוטציות המובילות וגילוי תזמון, פותחו גישות חדשות המנסות לגלות את הנתונים הללו בעזרת ניתוח פילוגנטי של תאים הלקוחים מהרקמה הנגועה, ובניית עצים מתוך האוכלוסייה. מתוך מבנה העץ, ניתן להסיק אודות זהות המוטציות עצמן, וכן הזמן היחסי בו הן התרחשו. במסגרת המחקר עליו נבסס את ההאקתון[1], זוהו 580,133 מוטציות סומטיות רלוונטיות, ומהן שוחזר עץ פילוגנטי. המחקר מצביע על כך שרכישת מוטציות מובילות מוקדמות והתפתחותן לכל אורך החיים הן הבסיס למחלות מיילופרליפרטיביות במבוגרים, דבר שמעלה הזדמנויות להתערבות מוקדמת ומציע מודל חדש להתפתחות סרטן.

### שאלות המחקר

\* כיצד מוטציות גנטיות מוקדמות ותהליכי התרחבות שיבוטיים משפיעים על התפתחותן של מחלות מיילופרליפרטיביות לאורך החיים?

\* באיזה שלב בחיים חולי סרטן מקבלים את מוטציות הדרייבר, המקדמות סרטן בהמשך החיים?

### מטרות הפרוייקט

\* נרצה לנצל את הידע הנלמד בקורס בנושא בניית עצים פילוגנטיים על מנת לזהות פיצולים (מוטציות) בהתפתחות התאים שמובילות לסרטן, ולהבין את התזמון של רכישת מוטציות מובילות.

\* שיחזור ואימות תוצאות המאמר עליו אנחנו מתבססים.

\* הכרת עבודה עם דאטה ביולוגי בפורמטים חדשים.

## דאטה

הקישור למקור הדאטה מצורף כאן.  
המידע נלקח מתוך קובץ נתונים גנומיים בפורמט VCF (Variant Call Format), שמטרתו לתאר מוטציות גנטיות באזורים ספציפיים בגנום. במקרה זה, מדובר במוטציות סומטיות שנמצאו בתאי דם שנדגמו מ-1013 קולונות, שמקורן ב-12 פציינטים. אלו מוטציות שנרכשו במהלך החיים ולא תורשתיות. ראו הרחבה על הדאטה ונספח ניתוח הפורמט.

## היפותזה

מוטציות מובילות שנרכשות בשלבים מוקדמים בחיים קובעות את הכניסה למצב המחלה בשלבים מאוחרים יותר בחיים, לפני האבחנה הקלינית. ממצא זה יוכל לשפר את הגילוי המוקדם ולתרום להתאמה אישית של טיפולים.

## שיטות חישוביות

- \* שימוש באלגוריתמים של בניית עצים פילוגנטיים (MPboot).
- \* שיטות לשחזור מבנה עצים פילוגנטיים - UPGMA, Neighbor Joining
- \* מודלים בייזיאניים לחישוב קצבי התפשטות קלונים.
- \* ניתוח סטטיסטי להשוואת מוטציות מניעות בין קלונים.

## תכנית עבודה

- \* ניתוח ועיבוד ראשוני של הנתונים ובדיקת איכות.
- \* בניית עץ פילוגנטי עבור הדאטה המצורף במאמר - בדיקת אופן בניית העץ (צימוד שכנים או אולטרמטרי), והשוואת גנים של חולים ברצפים הייעודיים לגנים בריאים ותיעוד היסטוריית המוטציות.
- \* ניתוח התוצאות מהעץ ומסקנות.

## עילה לבונוס

- הקבוצה שלנו כוללת אנשים מרקעים שונים ומגוונים:
- ✓ 3 סטודנטים לביולוגיה חישובית, סטודנט למדעי המחשב חד חוגי, סטודנטית למדעי המחשב עם חטיבה בביולוגיה.
- ✓ 3 עתודאים, 1 קצין במיל' (תותחנים), 1 קצינה בקבע (חיל הים) - שילוב הזרועות השונות.
- ✓ 2 דתיים ו-3 חילוניים.
- ✓ ג'ינג' אחד.
- ✓ גם בנים וגם בנות.
- ✓ 4 משקפופרים ואחד שרואה טוב.
- הקבוצה שלנו מגוונת ברקע האקדמי, המקצועי והאישי של חבריה. אנו כוללים סטודנטים ממסלולים שונים כמו ביולוגיה חישובית ומדעי המחשב, לצד עתודאים, קצין במילואים וקצינה בקבע מחילות שונים. בקבוצה יש ייצוג של דתיים וחילוניים, בנים ובנות, וכן גיוון אתני המשקף את החברה הישראלית. אלמנטים אישיים כמו ג'ינג' אחד ומשקפופרים מוסיפים צבע וגיוון נוסף. המגוון שלנו מאפשר שיתוף פעולה ייחודי וחשיבה רחבה ויצירתית.

**מקורות**

Williams, N., Lee, J., Mitchell, E., Moore, L., Baxter, E. J., Hewinson, J., Dawson, K. J., [1]  
Menzies, A., Godfrey, A. L., Green, A. R., Campbell, P. J., & Nangalia, J. (2022). Life  
histories of myeloproliferative neoplasms inferred from phylogenies. *Nature*, 602(7895),  
162–168. <https://doi.org/10.1038/s41586-021-04312-6>