



פרופ' איתן דומאני בעת ההרצאה



קשר לנושא לימוד:
התא, מידע
ותקשורת, עיבוד
מידע.

ניתן להשתמש במאמר
זה ובמאמרו של ד"ר
אלכס רזניק בגיליון
זה, "מדע חדש -
התחלה חדשה", "לשם
דיון בצמיחתם של
תחומים חדשים במדע
שנולדו מתוך חיבור
של תחומים קיימים.

תודה לגדי גץ על
עזרתו בהגהת המאמר

חקר
ביטוי
גנים
בעזרת
שבבי
DNA.

למה
עוסק בזה
פיסיקאי?

רשמה מירה משולם

על-פי הרצאה של פרופ' איתן דומאני, המחלקה לפיסיקה של מערכות מורכבות, מכון ויצמן למדע, שניתנה ביום פתיחת
המרכז הארצי למורי מו"ט בחט"ב אוקטובר 2000.

בהכשרתי אני פיסיקאי, אולם אני עוסק בבעיות בביולוגיה. מדוע? כי אני מאמין שב10-15 השנים הבאות יבואו פריצות הדרך
הקונצפטואליות העיקריות במדע מתחום הביולוגיה, ולא מהפיסיקה. לשואלים אותי מה כדאי ללמוד, אני עונה: כדאי ללמוד פיסיקה,
ואחר-כך לעבור לביולוגיה ...

לב השיטה שפיתחנו לחקר של ביטוי גנים היא מערכת פיסיקלית המשמשת אותנו לפתרון בעיות של חלוקת נתונים לצברים
(clustering of data). בעזרת חלוקה לצברים אנו מנתחים תוצאות של ניסויים המבוססים על הטכנולוגיה החדשה של שבבי
DNA



ואילו לזברות יש צוור קצר וכתמים ארוכים וצרים .

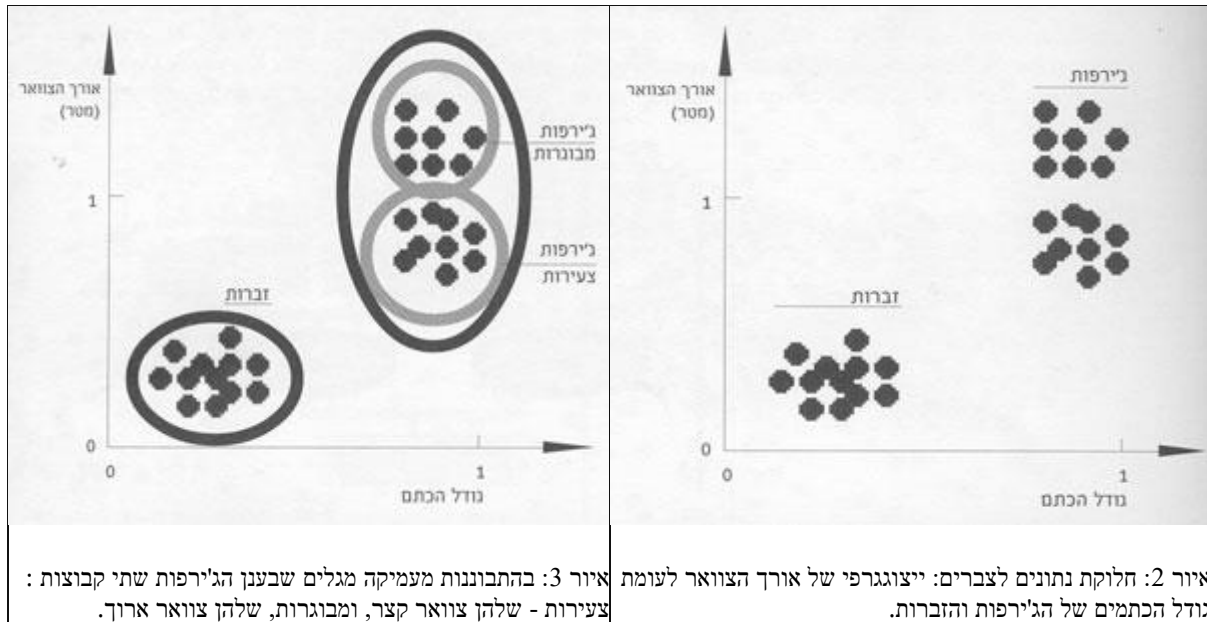


איור 1: לג'ירפות יש צוור ארוך וכתמים עגולים

מהי חלוקת נתונים לצברים?

העולם שמסביבנו מכיל מידע רב. אנו מעבדים אותו במוחנו כל הזמן, ואת רובו ללא עזרת מורה. אנו, המדענים, רוצים למעשה שתהיה לנו תוכנת מחשב, שתעשה דבר דומה.

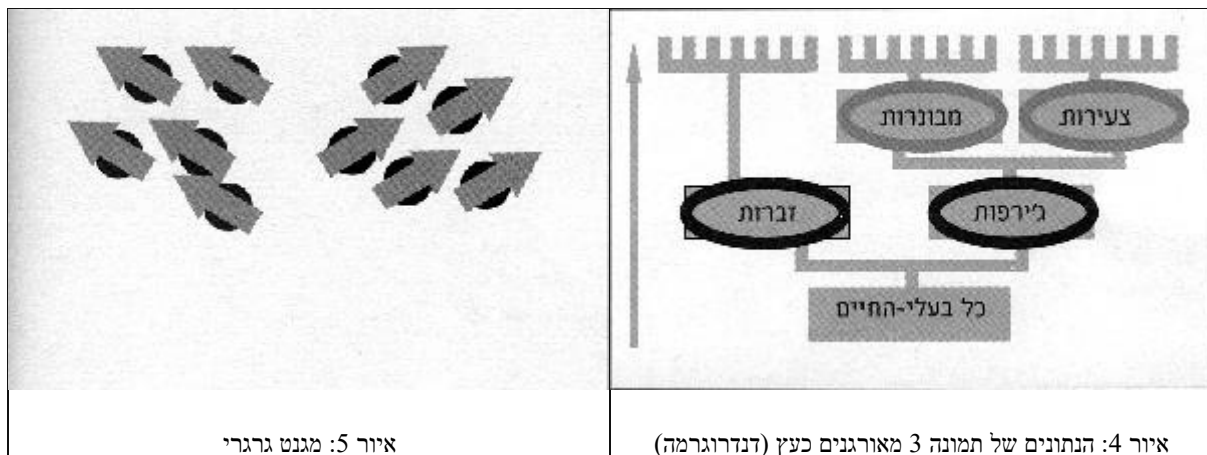
נתחיל בדוגמה. נחשוף ילד קטן לתמונות של ג'ירפות וזברות, אך לא נאמר לו דבר עליהן. הוא יתבונן בהן, יקלוט את המידע, ובאופן לא מודע יעשה מדידות והערכות שונות, כמו: כמה רגליים יש לבעל-החיים? מהו אורך הצוואר? מה צורת הכתמים שעל הגוף? מה גודלם? אם נראה לו תמונות רבות, יבין הילד שהוא רואה שני בעלי-חיים שונים. אומנם לשניהם יש ארבע רגליים, אולם לג'ירפות יש צוואר ארוך וכתמים עגולים ואילו לזברות יש צוואר קצר וכתמים ארוכים וצרים (איור 1). בעזרת נתונים אלה הוא יצליח להבחין בין שתי הקבוצות. למעשה הוא ביצע חלוקת נתונים לצברים.



ניתן להדגים זאת בייצוג גרפי של אורך הצוואר לעומת גודל הכתמים (למעשה הנתון הוא היחס שבין רוחב הכתם לאורכו) (איור 2). כל פרט, זברה או ג'ירפה, עובר בדיקה: האם יש לו צוואר ארוך או קצר? מה גודל הכתמים? כל פרט מבוטא בתוצאה על-ידי נקודה במישור, ולבסוף רואים שהנקודות מסודרות בשני "עננים": ה"ענן" האחד מייצג את הזברות, שלהן צוואר קצר וכתמים ארוכים, והשני את הג'ירפות, שלהן צוואר ארוך וכתמים עגולים.

אולם אם נתבונן קצת יותר "לעומק", נראה שבענן הג'ירפות יש שתי תת-קבוצות: הצעירות, שלהן צוואר קצר, והמבוגרות שלהן צוואר ארוך. מבחינת הכתמים אין הבדל גדול ביניהן (תמונה 3).

חלוקת הנתונים לצברים איננה בעיה המוגדרת היטב. צורת החלוקה, כלומר מספר הצברים, תלוי ברזולוציה (resolution) בדרגת "העומק" של ההתבוננות. ניתן לרשום את הנתונים גם כעץ (דנדרוגרמה): מתחילים מהענן הראשון, המייצג את כל בעלי-החיים, והוא מתחלק לענן של ג'ירפות וענן של זברות. ברזולוציה גדולה יותר, מתחלק ענן הג'ירפות לצעירות ולמבוגרות (איור 4).



איך פותרים את הבעיה של הצגת הנתונים?

הפתרון נבע מהסתכלות במערכת פיסיקלית, הנקראת "מגנט גרגרי" (granular magnet) הקיים בכל רשמקול. חומר זה עשוי גרגרים מגנטיים, וכל גרגר מגנטי עשוי ממונטים מגנטיים קטנים. ביניהם קיים צימוד חזק מאוד, כי הם קרובים מאוד זה לזה (איור 5). לעומת זאת, קיים צימוד חלש בין מונטים היושבים על גרגרים שונים. אם נסתכל על אטומים או יונים של המגנט הגרגרי, נראה שהמבנה של הנקודות במרחב דומה לזה שראינו עם הג'ירפות והזברות: ענן של אטומים מייצג גרגר 1, רחוק ממנו יש ענן 2, ואם

מסתכלים היטב, רואים שהוא בנוי משני תת-גרגרים. השתמשנו באנלוגיה זו בין מערכת מגנטית גרנולרית לבין המבנה הגיאומטרי של הנתונים שרצינו לסדר לצברים.

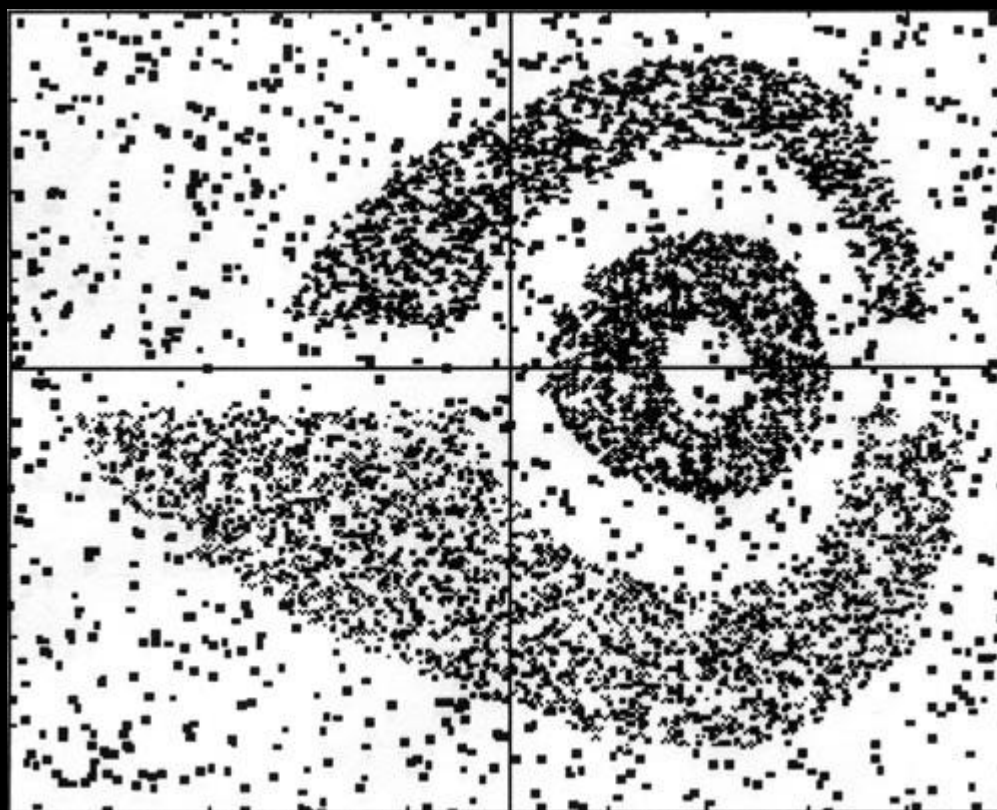
התבוננו באיור 6: יש בו כ-7,000 נקודות. כמה צברים אתם רואים? סביר להניח שתאמרו שיש כאן שלושה צברים גדולים - אבל יש גם צברים קטנים, ולא ברור איך לחלק אותם. לפיכך פיתחנו תוכנית מחשב שתקבע במדויק כמה צברים קיימים. כשיש קואורדינטות ונקודות במרחב, מחלק המחשב את הנקודות לקבוצות, וקיים סיכוי גבוה שנקודות הקרובות אלה לאלה תשויכנה לאותה קבוצה.

מטרות חקר התבטאות גנים

אנו יודעים אפוא לחלק נתונים לצברים בעזרת המחשב, ומשתמשים בשיטה זו לחקר ההתבטאות של גנים. המטרות העיקריות של חקר זה הן:

1. לזהות גנים המשתתפים בתהליכים ביולוגיים.
2. לפתח כלי לאבחון (דיאגנוסטיקה) ולריפוי.

למשל: הגנים האחראים לתהליך סרטני ברקמה מסוימת שונים מאלה המתבטאים ברקמה, שאינה סרטנית. לכן אם נצליח לזהות את הגנים המשתתפים בתהליך, יהיה לנו כלי דיאגנוסטי, ואולי גם דרך לריפוי. כלומר: כאשר מבינים את התהליך הגנטי המביא למחלה מסוימת, יש אולי אפשרות לחסום את המסלול המוביל למחלה או לנבא איזו תרופה תהיה מתאימה לחסום את התבטאות הגן. בנושא זה יש מחקר בסיסי ושאלות רבות שאינן פתורות.



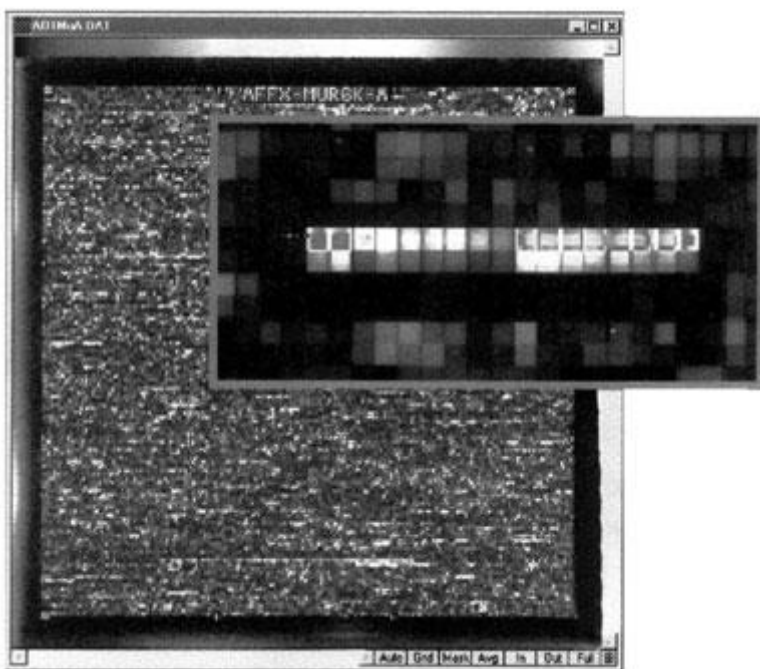
איור 6: כמה צברים אתם רואים בתמונה?

כאשר מבינים את התהליך הגנטי המביא למחלה מסוימת, יש אולי אפשרות לחסום את המסלול המוביל למחלה או לנבא איזו תרופה תהיה מתאימה לחסום את התבטאות הגן.

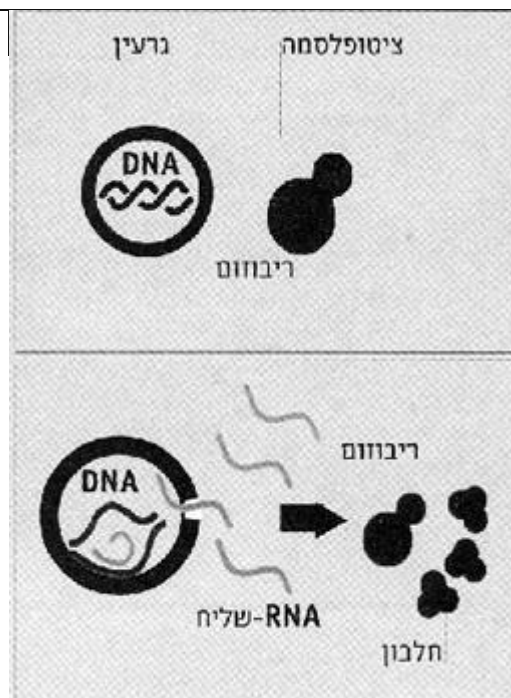
גילוי גנים של
סרטן התאפשר
הודות לשילוב
של שיטות
פיסיקליות,
ביולוגיות
וחישוביות.

תהליך ההתבטאות של הגנים בתאים

גן הוא קטע של DNA הנמצא בגרעין התא החי המכיל את התוכנית (המידע הנחוץ) ליצירת חלבון. כל גן אחראי על ייצור חלבון מסוים, כלומר יש בו מידע על הנוסחה הכימית של אותו חלבון. המפתיע הוא שכל אחד מהתאים שלנו מכיל בגרעין שלו את כל ה-DNA. כלומר את כל הגנים האנושי, המקודד 80 עד 100 אלף חלבונים שונים! אכן, כל תא מכיל את כל המידע, אולם לא כל תא מייצר את כל החלבונים הללו. למשל: תאי הכבד ותאי הרשתית אינם מייצרים את אותם החלבונים: תאי הכבד אינם מייצרים מולקולות חלבון הרגישות לאור (על אף שהמידע מצוי בהן), ובעין לא נוצרים חלבונים הפועלים את פעולת הכבד. יש אפוא חלוקת עבודה: כל תא מייצר חלבונים מסוימים, כלומר רק מיעוט קטן של גנים בא לידי ביטוי בכל תא. אפשר לומר שרוב הגנים "שקטים". וכאשר גן מתבטא, סוף תהליך ההתבטאות הוא ייצור (סינתזה) של חלבון שעליו הוא אחראי. ייצור זה מתבצע בבית-החרושת של התא, בריבוזום, הנמצא בציטופלסמה (איור 7). נשים לב: יש רק תוכנית אחת לסנתוז החלבון, והיא נמצאת בתוך הגרעין (המהווה העתק אחד "יקר"). אם התוכנית ניזוקה, היא תייצר חלבון לא נכון, שגוי, וזה יפגום בפעולה שאותו חלבון אמור לבצע. כאשר תוכנית הסנתוז ניזוקה, הדבר יתבטא במחלה. בנוסף: איך יעבור המידע מהגרעין לריבוזום? כדי להמחיש את הבעיה, חשבו על תוכנה חשובה שיש לכם על הדיסק הקשיח במחשב. אם תהיה לכם בעיה בדיסק, תאבדו את התוכנה; ואם תרצו להעביר את התוכנה למישהו אחר, ודאי שלא תעבירו אליו את הדיסק הקשיח כולו. מה עושים? מכינים דיסקט, ובו גיבוי של המידע, ואותו מעבירים לנמען. בתא מתרחש מהלך הדומה להכנת הדיסקט. בגרעין (באנלוגיה - הדיסק הקשיח) מיוצר החלבון RNA-שליח (Messenger RNA ובאנלוגיה - הדיסקט). הוא עובר מן הגרעין אל הריבוזום ("בית-החרושת" של התא) ומעביר לו את המידע המדויק של החלבון שעליו לסנתוז (איור 8). כל RNA-שליח מתאים לחלק של ה-DNA ומבטא את המידע שנמצא בו. כאשר גן מתבטא, הציטופלסמה של התא מוצפת במולקולות RNA-שליח. במילים אחרות: התוכנית לבניית חלבון מסוים נשלחת מהגרעין אל הריבוזום. כשרוצים לדעת ברגע מסוים אם גן מתבטא או לא, ניתן אפוא למדוד את ריכוז ה-RNA-שליח שלו בתוך הציטופלסמה של התא. ההתבטאות של הגנים ברגע נתון משקפת את המצב של התא. אנו מניחים, שאם יודעים את רמת ההתבטאות של כל אחד ואחד מהגנים ברגע נתון, די בכך כדי לקבוע אם מתרחש בתא תהליך בריא - או תהליך פתולוגי. לכן יש חשיבות גדולה למדידת המידע, שבה באים הגנים לידי ביטוי. אם גן מתבטא - יהיה ריכוז ה-RNA-שליח שלו גבוה. אם אינו מתבטא - יהיה הריכוז נמוך. שיטת המדידה של ריכוז RNA-שליח בתא היא באמצעות שבב של DNA (DNA chip, Gene chip). נתאר כיצד בנויים אותו.



איור 9: שבבי DNA בנקודות הבהירות נדבקה כמות גדולה של-RNA שליח.



למעלה: איור 7: סכימה של תא.
למטה: איור 8-RNA: שליח עובר מן הגרעין אל הריבוזום ומעביר לו את המידע המדויק של החלבון שעליו לסנתז.

מדידה בעזרת שבבי DNA

לוקחים שבב העשוי מסיליקון בגודל של כסנטימטר רבוע, המחולק לאזורים קטנים ("פיקסלים"). הוא מכיל כמה אלפי פיקסלים, ובכל אחד מהם שותלים חתיכת DNA מאותה רקמה, שבה רוצים למדוד את מידת ההתבטאות של הגנים. כל פיקסל מכיל אפוא חתיכה אחרת, כלומר גנים אחרים.
עתה מפרקים את הרקמה, עושים ממנה תמיסה, ושופכים אותה על השבב. ברקמה יש RNA-שליח - אך ורק של אותם גנים שבאו לידי ביטוי בתאים - והמולקולות של ה-RNA-שליח נדבקות ל-DNA וכל אחת לחלק המתאים לה, כלומר: רק לחלק מהפיקסלים יתקשרו מולקולות של RNA.
איך מגלים איפה יש התקשרויות?
משתמשים בטכנולוגיה מיוחדת: מדביקים למולקולות של ה-RNA סמן פלואורסצנטי (מעין "פנס") כדי שאפשר יהיה להבחין בהן. בסופו של התהליך, כשמתבוננים בשבב, רואים נקודות בהירות. אלה הפיקסלים, שבהם נדבקה כמות גדולה של RNA-שליח, והם בולטים בגלל החומר הפלואורסצנטי. (איור 9).
בסופו של הניסוי מקבלות הנקודות המוארות ביטוי מספרי: מקבלים אלפיים מספרים, המבטאים את רמת ההתבטאות של כל גן וגן שהושלל על השבב.

תיאור ניסוי: ביטוי גנים בסרטן

לקחו רקמות מ-62 אנשים: 40 רקמות שהיה בהן גידול, 221-רקמות נורמליות, ללא גידול, וביצעו את בדיקת ה-RNA-שליח בעזרת שבבים כמתואר קודם. על כל רקמה קיבלו 2,000 מספרים, ובסך-הכל 124,000 מספרים, המייצגים ביטוי של גנים. כדי לעבד בסיס נתונים גדול כל כך ולהציג את הנתונים, יש צורך בשיטות עיבוד מתוחכמות: יוצרים מטריצה של 2,000 שורות (גנים) מ-621 עמודות (כמספר הרקמות), ועליה מציגים את התוצאות, ולאחריה מגיע השלב של חלוקת הנתונים לקבוצות, לצברים. הייצוג שלהם הוא בדנדרוגרמה (איורים 10, 11).
מצאנו שהרקמות מתחלקות באופן ברור לשתי קבוצות. לאחר מכן התברר שהאחת מייצגת את הרקמות הסרטניות, והשנייה את הבריאות. כל עמודה מייצגת רקמה. אפשר להסתכל על שורה המייצגת גן ולבדוק את מידת ההתבטאות שלו ב-62 אנשים. האם יש גנים שפועלים יחד בצורה קוהרנטית (עקבית)? מצאנו שמבין 2,000 הגנים ישנן שתי קבוצות ברורות מאוד: יש קורלאציה

בין פעילות הגנים בכל קבוצה.
אנו מחפשים חתימה של תהליך סרטני. קבוצת הגנים של הסרטן אינה מרוכזת, אלא מפוזרת, ולכן צריך "לסנן" ולגלות רק אותם.
ואכן, הצלחנו להתמקד בקבוצה קטנה יחסית של 100 גנים (אחת מהקבוצות שמצאנו). אם "מקשיבים" רק להם, מקבלים חלוקה ברורה מאוד של רקמות ל"סרטניות" ול"לא סרטניות".
ביצענו ניסוי זה בחולי לאוקמיה (סרטן הדם), ובדרך זו הפרדנו את הגנים של המחלה לשתי תת-קבוצות, וכל אחת מהן קשורה לסוגי תאים שונים הקשורים במחלה.
לסיכום: גילוי גנים של סרטן התאפשר הודות לשילוב של שיטות פיסיקליות, ביולוגיות וחשוביות. קיימים כיום שטחי מחקר רבים, בהם נוצרים שילובים חדשים ומרתקים של תחומי הידע המסורתיים, המובילים לפריצות דרך חדשות. וכך אני חוזר לכותרת המאמר: איך אני, כפיסיקאי, הייתי לחוקר ביולוגיה נלהב...

לקריאה נוספת

טראובמן תמרה, השבב המחולל את המהפכה, הארץ, 27.7.2000

