# תרגיל בית 1 – UCSC Genome browser

הנחיות כלליות:

* בבקשה נסו לפתור את כל השאלות המופיעות בתרגיל.
* אתם מתבקשים לרשום פתרון מלא ולא רק פתרון סופי. המשמעות: שניתן לשחזר את הפתרון שלכם. ניתן לפרט בקווים כלליים באופן דומה לתרשים ההנחיות. הקו הכללי – ניתן לשחזר את התוצאות שלכם.
* בשלבים הדורשים להגיע לתצוגה ויזואלית (ה-UCSC), מאוד ממליץ לבצע צילומי מסך (ניתן להשתמש ב-windows snipping tool, <windows key>+<shift>+<S>).
* ניתן כמובן להיעזר אחד בשני, אך הכתיבה וצילומי המסך עצמאיים.
* במידה ואתם בוחרים לכתוב בכתב יד: אין בעיה (כל עוד הוא קריא). תצרפו שני קבצים: אחד צילום כתב היד והשני צילומי המסך. חד משמעית יש עדיפות לכתיבה במחשב.
* בונוס: לציון התרגיל הנוכחי במידה וטעיתם במקומות אחרים :) לא חייבים לבצע אותו.
* בונוס מתרגילי הכיתה =] : לציון התרגיל הנוכחי. תרגיל 1 – 2 נק'. שני תרגילים – 5 נק'. שימו לב שלא ניתן לעבור את ה-100 והנקודות אינם נשמרות ;) . התרגילים הרלוונטיים: תרגילי כיתה 1 ו-2.
* התרגילים לשבועיים והתרגילי כיתה ייבדקו לאחר תום ההגשה.

בהצלחה!

## USCS

הנחיות:

* עיקר העבודה היא ב-UCSC - <https://genome.ucsc.edu/index.html>
* אנו נעבוד הפעם עם hg38 (Human Genome 38).
* אנא הקפידו לרשום את השלבים שביצעתם על מנת להשיג את מטרתכם. מומלץ לצרף צילום מסך של התוצאה (במידת הצורך).
* אין צורך להסתבך! תשובה לא אמורה לקחת יותר מ-1-2 שורות לכל היות
* הערה כללית: אם אתם מסתבכים, נסו להגדיר את הערוץ, חפשו מצד שמאל בשמות המרכיבים, בחרו תצוגת full ולא pack ולחצו על אלמנטים שונים במידה וצריך. אל תחששו להגדיל אזורים שונים שמעניינים אותכם וכמו כן – פשוט להעביר את העכבר עליהם, זה נותן כמעט את כל המידע.

### שאלה 1 – פרטים כלליים

חפשו את הגן MDM2:

1. כמה **refseq genes** קיימים לגן?
2. כמה GENCODE v38 genes יש?
3. האם הגן מקודד מהגדיל החיובי או השלילי? הסבירו את בחירתכם.

הסתכלו על האיזופרם בעל כמות האקסונים הרבה ביותר, וה-3’UTR הארוך ביותר ( ב- NCBI refseq אם לא פתוח. רמז: זכרו איפה ה-UTR נמצא).

1. כמה אקסונים/אינטרונים יש לגן?
2. האם כל האקסונים מקודדים לחלבונים? הביטו בשני האקסונים הראשונים). (רמז: zoom על האקסונים)?
3. כתבו את המיקום של הגן בפורמט BED (רמז: שורה אחת, הקפידו על פורמט תקני).
4. מה אורך החלבון הנוצר ממנו (הדגשה: החלבון, לא כמות הבסיסים! רמז: חשבו כיצד ניתן לקבל מידע על החלבון).

בונוס ( 1 נק'): האם מצאתם איזופורמים בגן? במידה וכן – שערו ממה נובעים האיזופורמים במקרה זה.

### שאלה 2: רמות ביטוי

באיזה רקמות הגן מתבטא בעיקר? כתבו את 2-3 סוגי הרקמות העיקריות עם הביטוי הרב ביותר (רמז: ערוץ **GTEx gene**).

בונוס (2 נק'): חוקר הפיק דגימת RNA מלימפוציטים ומדם וגילה שהגן MDM2 משנה את הביטוי שלו במספר מחלות אוטואימוניות. אחת הביקורות הינה האם אכן יש ביטוי של MDM2 ברקמת הדם (whole blood) ובלימפוציטים? הכנסו לפורטל ה-GTEx ובדקות את רמות ה-TPM בשני הרקמות (רמז: יש לכם קישור ישיר מה-UCSC).

### שאלה 3: ביטוי mRNA

השתמשו בערוץ **Human mRNAs**.

כמה mRNAs ידוע שמקורם מרקמת הלימפוציטים (**‘lymphocytes’**)? רשמו את שמותיהם.

הערה: שימו לב שעליכם לחפש באופן כללי בכלל רקמות הלימפוציטים, התעלמו משאר הרקמות. הפתרון לשאלה זו יותר קל מהמבט הראשון. חשבו, כיצד (ואיפה) ניתן לסנן רקמות, וככלל היכן ניתן להגדיר פרטים בערוץ. כמובן, הראו דרך מליאה.

### שאלה 4: אלמנטים חוזרניים

השתמשו בערוץ **RepeatMasker**:

כמה אלמנטים מסוג SINE הנכם רואים באינטרון העשירי? רשמו את המשפחה ואת המיקום שלהם.

בונוס (2 נק'): אלמנטי ALU (ממשפחת ה-SINE) בעלי חשיבות מיוחדת היות והם יכולים להכיל אתרים מקודדים (או כמעט מקודדים). הם נוטים להתחבא באינטרונים/UTR כשלב ראשון, ואז מוטציה בהם יכולה ליצור שינוי בכל הגן.

הביטו על האיזופרמים של הגן ועל האלמנטים החוזרניים. האם יש הבדל באורך ה-3’UTR? האם יש אלמנטים חוזרניים שנוטים להופיע שם? במידה וכן, רשמו את שם המשפחה וצרפו צילום מסך המראה זאת.

### שאלה 5: SNP

השתמשו ב**Common SNPs(153)**. מצאו את מוטציית ה-synonymous המצויה שם? (רמז: חשבו על הצבע ☺, אתם לא אמורים לראות יותר משתיים).

רשמו את שם ה-SNP ואת המיקום שלו.מה שכיחות האללים באוכלסיה?

### שאלה 6: CRISPR

שיטת קריספר (CRISPR) הינה שיטה גנומית מהפכנית לעריכת דנ"א, אשר בעלת הרבה פוטנציאל. מאז המצאת השיטה היא נכנסה לשימוש כמעט רגיל במעבדות רבות – וכעת גם שני חוקרים על הגילוי פרס נובל.

1. פתחו את הערוץ **CRISPR Targets** והציגו אותו. התמקדו באקסון האחרון/ב-3’UTR. בחרו אתר אחד של CRISPR לחיתוך אשר בעל ציון גבוה לספיציפיות (ניתן לקרא על הגדרת ציון גבוה בעמוד ההגדרות של הגן) (רמז: השתמשו בתצוגה full והעבירו את העכבר על האתרים השונים. שימו לב גם לצבעים).
   1. מה הציון של הספיציפיות? (MIT Guide Specify).
   2. מה המיקום של האזור?
   3. האם יש off-targets ל-Guide? תנו דוגמה להם (ניתן לצלם מסך).

הערה: אל תסתבכו עם השאלה! כל המידע מופיע לכם במקום אחד ובדף אחד. כמובן יש יותר מתשובה אחת נכונה.

### שאלה 7: שמירות הגן

הסתכלו בערוץ **Conservation** ב-Comperative genomics:

1. הביטו בגן ככלל. אילו אזורים שמורים יותר? אקסונים/אינטרונים?

חוקר מעוניין לבדוק את השפעתה של תרופה חדשה שנועדה לווסת את רמות MDM2. התרופה נקשרת אל ה-3’UTR ומווסתת את הביטוי של החלבון הנוצר. החוקר מעוניין לבדוק את התרופה קודם כל על עכברים ואז להמשיך לבדוק את התרופה על אורגניזמים אחרים.

1. הביטו בשמירות של ה-3’UTR אצל עכברים (Mouse), האם האזור שמור היטב בעכברים?

רמז: הביטו בתת הערוץ Multiz Aligments of 100 vertabrates וסננו את האורגניזמים.

בונוס (1 נק'): במידה ועניתם שהאזור לא שמור אצל עכברים, האם תוכלו להציע אורגניזם אחר שיתאים למחקר?

### שאלה 8: הרצה ב-BLAT

כעת ניקח את רצף ה-DNA של הגן על מנת להשתמש בו באמצעות BLAT. הוציאו את רצף ה- mRNA של הגן שעבדנו עליו עד כה.

הדרכה: אין צורך להסתבך: לחצו על הגן בצד שמאל על מנת להגיע לדף המידע עליו, חפשו את ה-‘predicted mRNA’. שימו לב שלא השתמשנו הפעם בכלי ה-DNA.

פתחו כעת את הכלי של BLAT והעתיקו אליו את כל הרצף. בחרו את האורגניזם Gorila באנוטציה gorGor5. הפעילו את הכלי. הביטו על התוצאה עם ההתאמה הטובה ביותר.

1. רשמו את הציון ואת אחוז הזהות.
2. רשמו את התוצאה בפורמט BED

כעת חזרו על התהליך, אך הפעם חפשו כנגד Mouse באנוטציה mm10:

1. רשמו את הציון ואת אחוז הזהות.
2. רשמו את התוצאה בפורמט BED
3. הכנסו ל-details ובחרו מוטציה אחת של mismatch והציגו אותה בתמונה. ציינו את מספר הבלוק בו בחרתם להסתכל.

הדרכה: ניתן לבחור בכל מוטציה לבחירתכם. הסתכלו על התצוגה של שני האלמנטים אחד מעל השני ליתר נוחות.

בהצלחה מרובה!!!