המחלקה להנדסת תוכנה

פרויקט גמר – תשע"ו

בנאי פריימרים ל-DNA

DNA PRIMER BUILDER

מאת:

אמיר ארז - 203862248

אור ברדוגו - 302891551

מנחה אקדמי: ד"ר יהודה חסין **אישור:** **תאריך:**

רכז הפרויקטים: מר אסף שפיינר **אישור:** **תאריך:**

#	מערכת	מיקום
1	מאגר קוד	github.com/erezam/DNA-PRIMER-BUILDER
2	יומן	https://github.com/erezam/DNA-PRIMER-BUILDER/wiki/Project-Diary https://calendar.google.com/calendar/embed?src=ofbe3kdsde039ra5vqdq53tm9s%40group.calendar.google.com&ctz=Asia%2FJerusalem
3	ניהול פרויקט (אם בשימוש)	
4	הפצה	

תוכן עניינים

מילון מונחים: 3
מבוא 4
תיאור הבעיה 7
דרישות ואפיון הבעיה 8
הבעיה מבחינת הנדסת תוכנה 8
תיאור הפתרון 9
תיאור הפתרון המוצע 11
תיאור הכלים המשמשים לפתרון 11
סקירת עבודות דומות בספרות והשוואה 12
תכנית בדיקות 16
גרסת אלפא: 19
נספחים 22
רשימת ספרות \ ביבליוגרפיה 22
תכנון הפרויקט 22
טבלת סיכונים 23
טבלת דרישות 24
תרשים תפקידים כללי : 25
Sequence diagram 26
Class diagram 27
ERD 28

מילון מונחים:

- DNA - מולקולת ענק המורכבת ממספר רב של נוקליאוטידים המאורגנים במבנה של סליל כפול. כל המידע התורשתי הדרוש לבניית החלבונים בתא אצל כל האורגניזמים הידועים, מחיידקים ועד לבני אדם, מוצפן באחת או יותר מולקולות DNA שרצף הנוקליאוטידים בהן מיוחד לכל אורגניזם.
- RNA -
- נוקליאוטידים - קבוצה של תרכובות אורגניות. הנוקלאוטידים נמצאים בכל היצורים החיים, ומרכיבים, בין השאר, את החומר התורשתי. מיוצגים באותיות T,A,G,C .
- פריימר - מקטע קצר של חומצת גרעין המסייע בתהליך שכפול ה-DNA.
- מוטציה - הוא מונח בביולוגיה המתאר שינוי ברצף הנוקלאוטידים ב-DNA, דהיינו, שינוי בהרכב הגנטי של יצור מסוים. שינוי כזה עשוי להביא לשינוי באחת או יותר מתכונות היצור החי.
- PCR - היא שיטה מעבדתית המשמשת לשכפול מזורז של מקטעי DNA.
נהוג להשתמש ב-PCR לסריקה אחר מחלות גנטיות ולביצוע ניתוח השוואתי של DNA מאוכלוסיות שונות.
- אקסון – רצף נוקלאוטידים ב-DNA המהווה חלק מגן ואשר מקודד לחומצות אמינו. שאר חלקי הגן אשר אינם מקודדים לייצור חלבון קרויים אינטרונים.
- טרנסקריפט – Transcript - תהליך בביולוגיה של התא שבו מולקולת RNA מיוצרת לפי תבנית של מולקולת DNA , חיבור כל האקסונים בגן מסוים.
- Amplicon – אמפליקון – אמפליקון הוא חתיכת DNA או RNA המהווה מקור ו/או תוצר של צירוף מלאכותי או אירועים חוזרים. ניתן להפקה באמצעות מגוון שיטות לרבות תגובות שרשרת פולימריות או שכפול גנים טבעי.
אמפליקון הוא למעשה תפוקת של העתק אחד או יותר של חלק גנטי שמבוצע בו שימוש, בעיקר בתנאי מעבדה, כתוצר ממכשיר PCR.

מבוא

הפרויקט מבוצע כפרויקט מחקרי בשיתוף עם ד"ר מוטי בן שבת מבית החולים שיבא.

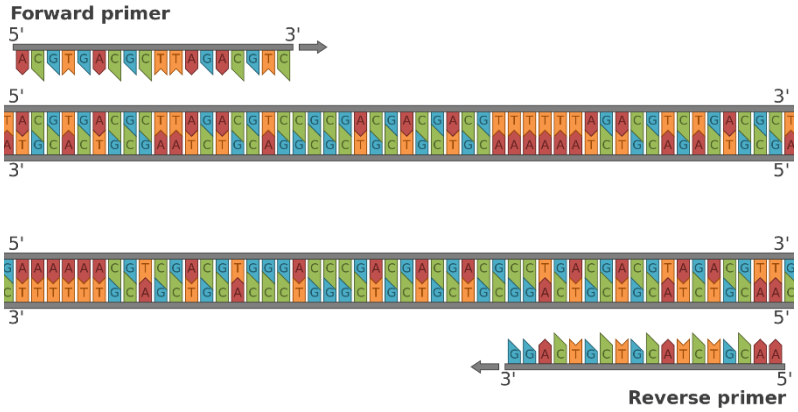
התוכנה שנבנה בפרויקט הינה תוכנה לעיצוב פריימרים שישמשו לבדיקות RT-PCR מהסוג של חיפוש גן מסוים בDNA.

מהו פריימר ?
פריימר (באנגלית: Primer; בתעתיק לעברית לעתים, במיוחד בעגה המדעית: פְּרָיימר) הוא מקטע קצר של חומצת גרעין המסייע בתהליך שכפול ה-DNA.
אורכו של התחל הוא כ-20 נוקלאוטידים בערך. הסיבה לקיומו היא אי-יכולתו של האנזים DNA פולימראז להתחיל ביצירת גדיל חדש של DNA. האנזים, האחראי על שכפול DNA בתא, מסוגל אך ורק להאריך גדיל DNA קיים. התחל מספק לאנזים נקודת התחלה ליצירת גדיל חדש.
פריימרים מיוצרים באופן מלאכותי על ידי חברות ביוטכנולוגיה ומשמשים רבות בשיטות בגנטיקה ובביולוגיה מולקולרית, כדוגמת PCR.

תהליך העבודה עם פריימר :

1. נסתכל על מקטע הDNA הנבדק,
קטעי DNA מורכבים משתי שרשראות הקשורות זו לזו (אנו נתייחס אליהן כשרשרת עליונה ותחתונה לשם הפשטות) , כל שרשרת מורכבת מרצף של נוקלאוטידים A,C, G ו-T
כאשר הקשרים בין השרשראות הם מהסוג T-A או C-G .
2. נסתכל על כל שרשרת בחלק נפרד , הסימונים '3 ו – '5 מסמנים בעצם את כיוון העתקת הרצף בגוף (מ5 ל-3). אנו צריכים לייצר שני פריימרים (ההסבר לסיבה בהסבר על הPCR) .
נבחר את הקטע המתאים ברצף DNA הנבדק בהתאם לאילוצים שנתאר בהמשך ונתחום אותו.
3. אחרי שתחמנו את הקטע המתאים נוכל לבנות את שני הפריימרים:
   - FORWARD – יבנה מתחילת הקטע התחום בכיוון 5 ל-3 ויהיה בעצם תעתיק של קטע מהשרשרת העליונה.
   - REVERSE – יבנה לפי סוף הקטע התחום ויהיה תעתיק של השרשרת התחתונה.

המטרה היא שהפריימרים ידבקו למקום התואם להם במקטע DNA ובכך יגרמו לשכפולו (הסבר בהמשך).

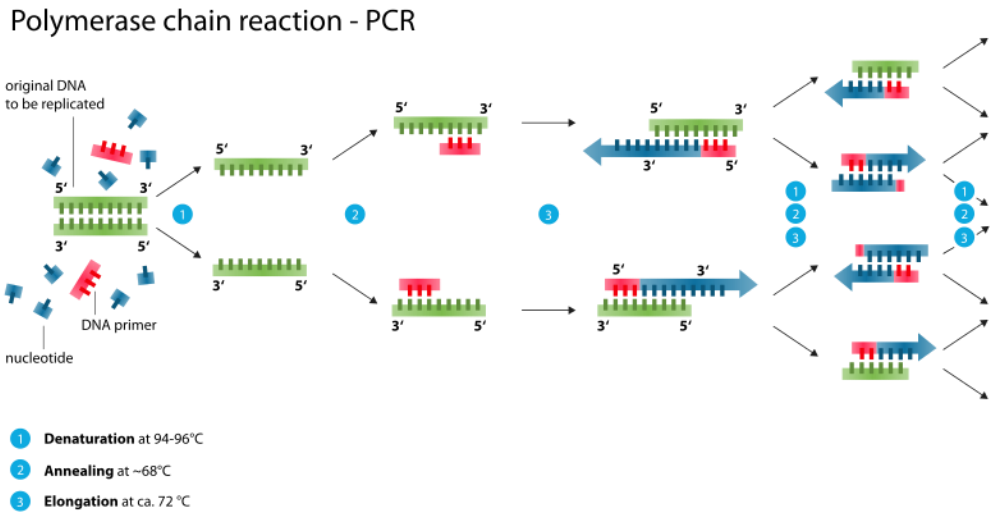


PCR היא שיטה מעבדתית המשמשת לשכפול מזורז של מקטעי DNA.
נהוג להשתמש ב-PCR לסריקה אחר מחלות גנטיות ולביצוע ניתוח השוואתי של DNA מאוכלוסיות שונות, כולל DNA ממינים נכחדים. השיטה חיונית לפענוח פשעים ולקביעת אבהות.

תהליך ה-PCR:

1. פרימה - מפרידים את שרשראות הDNA באמצעות חימום (מעל 90 מעלות).
2. איחוי - מנמיכים את הטמפ' עד לטמפ' המתאימה לפריימרים (חישוב נוסחתי), פריימר הפורוורד נקשר לתחילת הקטע הנבדק בשרשרת התחתונה , ופריימר הרברס לסוף הקטע הנבדק בשרשרת העליונה.
3. תחלים - תהליך השכפול מתחיל בכל שרשרת בכיוון ה'5 ל- '3 (שינוי טמפ' לתנאי שכפול)
4. חוזרים על שלבים אלו, של מחזור ההכפלה, כמה עשרות פעמים (בדרך כלל 25-30), עד להגברה המבוקשת של המקטע.

בהנחה של יעילות מקסימלית, בכל מחזור מוכפל מספר העותקים במבחנה פי 2. לכן, אם מתחילים עם עותק אחד, הוא ייתן בסוף המחזור הראשון שני העתקים, בסוף המחזור השני ארבעה העתקים, בסוף המחזור השלישי יהיו שמונה, וכך הלאה. הגברה כזאת מכונה הגברה מעריכית.
זה גידול מהיר מאד, לדוגמה, לאחר 30 סיבובים, מספר הגנים במבחנה מוגבר פי יותר ממיליארד.
התוצר של ה-PCR שמכונה אמפליקון הוא גן מוגבר פי מיליארדים, שיכול לשמש לליגציה או לקביעת רצף של גן.
לדוגמא: נניח לבדוק האם גן מסוים נמצא בDNA של אדם נבדק, נבצע את התהליך על דגימת הDNA של הנבדק בתוספת הפריימרים.
אם בתום התהליך נראה שכפול גדול מאוד של הגן , נדע שהוא קיים אצל הנבדק.



לסיכום :

התוכנה שנבנה תעבוד על טרנסקריפט (S) עבור גן מסוים , שהוא בעצם סטרינג באורך של סביבות 1000 תווים המורכב מחיבור של מספר אקסונים (רצף נוקלאוטידים ב-DNA המהווה חלק מגן ואשר מקודד לחומצות אמינו), כל נקודת חיבור בין שני אקסונים תקרא צומת.

נרצה למצוא שני פריימרים , פורוורד (F) ורוורס (R) סטרינגים באורך של כ-20 תווים בערך שיושבים על הצמתים של הטרנסקריפט, שנמצאים על אותו אמפליקון(A) בגודל של כ-80 עד 150 תווים (ז"א במרחק של 80-150 תווים)אחד מהשני ועומדים באילוצים עליהם נפרט בהמשך.

כל האותיות – טרנסקריפט S.

אדום – F - Forward primer

ירוק – A - Amplicon

סגול – צמתים – נקודות החיבור של האקסונים

תכלת – R - Reverse primer

*בהנחה שהפריימרים עומדים באילוצים.

```
GGGCTTGTGGCGCGAGCTTCTGAAACTAGCCGCCAGAGGCGGAGCCGCTGTGGCACTGCT
GCGCCTCTGCTGCGCCTCGGGTGTCTTTTGCGGCGGTGGGTCGCCGCCGGGAGAAGCGTG
AGGGGACAGATTTGTGACCGCCGCGGTTTTTGTCAGCTTACTCCGGCCAAAAAAGAACTG
CACCTCTGGAGCGGACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAAATGCCTA
TTGGATCCAAAGAGAGGCCAACATTTTTTTGAAATTTTTAAGACACGCTGCAACAAAGCAG
ATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAGAAGCTCCACCCTATA
ATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATTACGAACCAAACCTATTTA
AAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTTCAACTCCAATAATATTCAAAG
AGCAAGGGCTGAGTCTGCCGCTGTACCAATCTCCTGTAAAAGAATTAGATAAATTCAAAT
TAGACTTAGGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACAGTGAAAACTA
AAATGGATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGTCTTAGTGAAAGTC
CTGTTGTTCTACAATGTACACATGTAACACCACAAAGAGATAAGTCAGTGGTATGTGGGA
GTTTGTTTCATACACCAAAGTTTGTGAAGGGTCGTCAGACACCAAAACATATTTCTGAAA
GTCTAGGAGCTGAGGTGGATCCTGATATGTCTTGGTCAAGTTCTTTAGCTACACCACCCA
CCCTTAGTTCTACTGTGCTCATAGTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTC
ATGATACTACTGCTAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAAA
ATGATAGATTTATCGCTTCTGTGACAGACAGTGAAAACACAAATCAAAGAGAAGCTGCAA
GTCATGGATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATAGCTGCAAAGACCACA
TTGGAAAGTCAATGCCAAATGTCCTAGAAGATGAAGTATATGAAACAGTTGTAGATACCT
CTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTAGAACAAAAAATCTACAAAAAG
TAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATGAAGCAAACGCTGATGAATGTGAAA
AATCTAAAAACCAAGTGAAAGAAAAATACTCATTTGTATCTGAAGTGGAACCAAATGATA
```

תיאור הבעיה

הבעיה הקיימת הינה בניית פריימר מתאים עבור בדיקה גנטית במכונת PCR .

בבניית פריימר יש שני אספקטים מרכזיים:

- יעילות – הפריימר צריך לעמוד במבחנים בכדי לגרום לתהליך PCR תקין.
- נכונות – הפריימר צריך לעמוד במבחנים בכדי להתחבר למקטע הנכון.

הסביבה בה קיימת הבעיה היא בתחום המחקר והרפואה.

תיאור הליך בניית הפריימר :

בשלב הראשון צריכים לקבל את שם הגן שאנו מחפשים, נרצה לעבוד על הטרנסקירפט(ראה מילון מושגים) המרכזי של הגן בכדי להתרכז בחלקים החשובים בו ובחלוקה לאקסונים.

למעשה יש לבנות זוג פריימרים, forward primer ו reverse primer בהם משתמשים בבדיקה כפי שהוסבר במבוא.
לצורך מציאת פריימר מתאים לגן יש צורך לעמוד באילוצים הבאים:

1. אורך הפריימר - צריך לנוע בטווחים של 13-26 אותיות, כאשר 20 אותיות הוא האורך האופטימלי.
2. אורך האמפליקון – ינוע בין 80-150 תווים.
3. אחוז GC - (G,C הם 2 נקלאוטידים) צריך לנוע בטווח של 40%-60%.
4. ג'אנקשן אקסון – הטרנסקריפט מכיל אקסונים אחד אחרי השני, אנו מחפשים את נקודות המעבר בין אקסון לאקסון(junctions) ורוצים שהפריימר יתחלק באופן בו בין 40 ל 60 אחוז ממנו יהיו חלק מאקסון אחד והיתר יהיה באקסון אחר, ז"א שהפריימר יישב על נקודת החיבור.
5. טמפרטורה Tm - טמפרטורת פרימת הפריימר צריכה לנוע בין 60-70 מעלות צלזיוס, כאשר הטמפרטורה האופטימלית היא 65.
6. אסור שיהיה הבדל הגדול מ 5 מעלות בין ה forward primer לבין ה reverse primer.
7. אילוצים סינטקטיים:
   - הפריימר לא יתחיל בנוקלאוטיד G.
   - אסור שהפריימר יכיל רצף של 4 G ויותר.
   - אסור שהפריימר יכיל רצף של 6 A ויותר.
   - אסור שהפריימר יסתיים ברצף GGG או GGAG.
   - אסור רצף של CC.

תנאים אלו נלקחו ממדריך של תוכנת Primer Express שהיא התוכנה שמגיעה עם מכשיר ה PCR, יש לציין כי מקריאה מעמיקה באתרים שונים ובמאמרים שונים טווחי הפרמטרים משתנים ממקור למקור כך שהתנאים הנ"ל אינם סופיים כרגע.

דרישות ואפיון הבעיה

כאשר רופא או חוקר מעוניין לבצע שימוש במכונת PCR עליו למצוא פריימר מתאים לבדיקה.
כיום ישנן מספר דרכים למציאת הפריימר המתאים:

- מנועי חיפוש לפריימרים ידועים , עבור גנים שהחיפוש שלהם שכיח קיימים פריימרים תקינים וידועים.
- תוכנות לבניית פריימרים , תוכנות לבניית פריימרים בהתאם לגן ודרישות מסוימות.
- בניה ידנית של פריימר.

התהליך למציאת פריימר שעובד עשוי להיות ארוך ומתיש ,כל פריימר שחוקר בונה מוזמן מחברות שמייצרות את הפריימר ולאחר מכן עובר מספר ניסויים בכדי לבדוק האם הוא מתאים.
תהליך כזה לוקח כמספר חודשים, ולפי דברי הביולוג בהרבה מאוד מקרים הפריימרים שמתקבלים אינם עובדים.
ככל שלפריימר שהוזמן תהיה נכונות גבוהה יותר כך הסיכוי שהוא יעבוד עולה והתהליך מתקצר.

התוכנה שלנו תכיל מסך ראשי שבו תופיע תיבת טקסט להכנסת שם הגן שעבורו רוצים ליצור את סט הפריימרים.

המסך יכיל כפתור אישור להפעלת האלגוריתם והפקת סט הפריימרים.

לאחר שתהליך יצירת הפריימר יושלם יופיע הדף הבא שבו יופיעו הסטים של הפריימרים, הסטים יסודרו בסדר יורד ע"פ הציון, בכל סט יופיע דוח שבו רשומים האילוצים והערך של כל אילוץ.

המשתמש יוכל ללחוץ על כפתור שנמצא ליד כל סט פריימרים ולהוריד את הסט כקובץ למחשב.

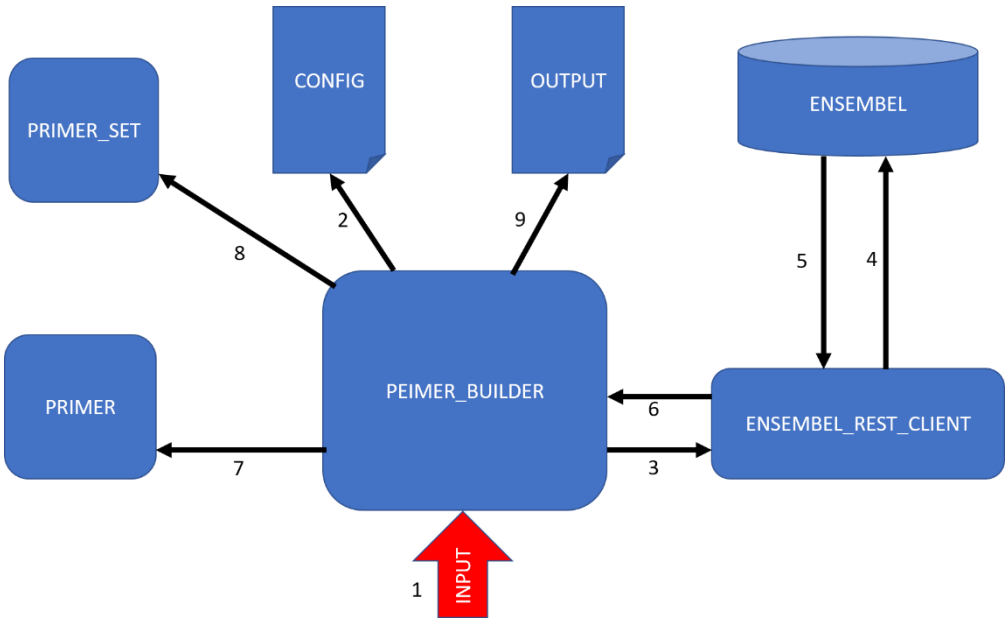
הבעיה מבחינת הנדסת תוכנה

מבחינת הנדסת תוכנה,
אפשר להתייחס למקטעים ב DNA(בין השאר גנים) כסטרינג(String).
הבעיה התוכנתית:

1. חיפוש הפריימרים האפשריים – התוכנה צריכה למצוא את כל זוגות הפריימרים(forward ו reverse) שיושבים כל אחד על חיבור של אקסונים ונמצאים במרחק תקין של אמפליקון.
2. בדיקת הנכונות – בדיקת התכונות של הפריימרים שנמצאו, האם הם עומדים בדרישות שהוזכרו בתיאור הבעיה (אחוז GC, טמפרטורה, אילוצים סינטקטים וכו'), טווחי הפרמטרים אמורים להיות דינמיים ולהיות מוגדרים בקובץ קונפיגורציה אותו יכול לערוך המשתמש כדי לשנות את הטווחים, דבר זה אמור לעזור לתוכנה להיות גנרית.
כרגע אנו נעבוד ע"פ הנתונים של תוכנה ליצירת פריימרים שנקראת primer express משום שהתכונות הן כללי אצבע שהתפתחו עם השנים ולא דבר קבוע ומוכח(וזו חלק מהבעיה).
3. נתינת ציון לכל תכונה, ובסופו של דבר ציון כללי לזוג פריימרים, ע"פ הציון הפריימרים יוצגו – מהטוב ביותר ומטה.
4. חיבור למסד נתונים חיצוניים בכדי למשוך מידע על רצפי הDNA ועל מוטציות מוכרות, יש צורך לבדוק כיצד להתחבר בעזר ה API של האתרים, ולשלוף את המידע הרצוי.

תיאור הפתרון

הפתרון ימומש ע"י תוכנה את שם הגן ואת שם הזן(אדם,עכבר וכד') ותחזיר את הפריימרים הטובים ביותר, בפרק זה נסביר על תהליך העבודה של התוכנה.



המודולים:

CONFIG – קובץ הקונפיגורוציות, קובץ המכיל את הנתונים על טווחי הפרמטרים לאילוצים, ניתן לשנוי לפי צורך החוקר, משמש את כל המחלקות בתוכנה, מאפשר גנריות ושינוי פרמטרים נוח ופשוט.

ENSEMBEL REST CLIENT – אחראי על משיכת הנתונים מהAPI של אתר ENSEMBEL, מושך קובץ JSON המכיל את כל הנתונים על הגן, ממנו נמשוך את הסטרינג של הטרנסקריפט, את מיקומי צמתי האקסונים ונתונים נדרשים נוספים.

PRIMER – מחלקה המייצגת פריימר בודד עם כל תכונותיו כגון: טמפ', אחוז GC אורך וכד'.

PRIMER_SET – מחלקה המייצגת זוגות פריימרים (פורוורד ורברס), מכילה את התכונות של כל זוג ואת הציון הסופי.

PRIMER_BUILDER – המודול המרכזי, מבצע את תהליך חיפוש ויצירת הפריימרים האפשריים.

OUTPUT – קובץ המכיל את הפריימרים שהופקו בתהליך , מסודרים ע"פ הציון מהגבוה לנמוך.

תהליך הריצה (לפי המספור בתרשים הנ"ל) :

1. המשתמש מזין לתוכנה את שם הגן והזן (לדוגמא :BRAF, human)
2. ה-MAIN מורץ מהPRIMER_BUILDER וקורא את הנתונים מקובץ ה-CONFIG
3. ה-PRIMER_BUILDER פונה לENSEMBEL REST CLIENT עם שם הגן והזן.
4. ה ENSEMBEL REST CLIENT שולח בקשה לשרת של ENSEMBLE לקבלת אובייקט ה JSON של הגן.
5. השרת של ENSEMBLE שולח ל ENSEMBEL REST CLIENT את אובייקט הJSON.
6. ENSEMBEL REST CLIENT שולח ל PRIMER_BUILDER את אוביקט הJSON.
7. PRIMER_BUILDER מנתח את אובייקט הJSON ומוציא ממנו את הסטרינג של הטרנסקריפט ואת מיקומי הצמתים של האקסונים, לאחר מכן יוצר אובייקטים של פריימר לכל פריימר אפשרי (פריימר שיושב על צומת של האקסונים ובטווחי האורך שמוגדרים בקובץ הקונפיגורציה) ומבצע עבור כל פריימר בדיקות האם הוא עומד באילוצים(GC%, טמפרטורה וכו') בסופו של דבר נשמרים הפריימרים התקינים .
8. לאחר מכן מתבצע חיפוש של זוגות תואמים לפריימרים שנמצאו, החיפוש מתבצע כך שהפריימר שמחפשים יהיה על אותו אמפליקון, הפריימר שיימצא ייבדק אם הוא עומד בכל האילוצים, לאחר מכן ייווצר אובייקט PRIMER SET המכיל את 2 הפריימרים שנמצאו, לאחר שיצר את כל הזוגות האפשריים העומדים בכל האילוצים מתבצע תהליך הענקת ציון לתכונות ובסופו של דבר ציון לכל פריימר ולכל אובייקט PRIMER SET.
9. PRIMER_BUILDER מייצא את כל אובייקטי ה PRIMER SET שנמצאו מסודרים ע"פ הציון.

* לתרשימים נוספים ראה נספחים.

תיאור הפתרון המוצע

הפתרון המוצע יכלול אתר אינטרנט שיהיה נגיש לכל אחד, באתר יוזן שם של גן לפי קונבנציית השמות של אתר Ensemble ושם הזן (אדם, עכבר, דג וכו').

לאחר הזנת השם התוכנה תשלוף בעזרת API את פרטי הגן כאובייקט JSON מהאתר ensembl,

התוכנה תבצע עיבוד של הנתונים ויצירת הפריימר ע"פ אלגוריתם שנבנה – האלגוריתם ייבנה לענות על האילוצים שהוזכרו בתיאור הבעיה וכמו כן ימתח את האילוצים ככל הניתן בכדי למצוא פריימר ללא התערבות ידנית.

האלגוריתם אמור להיות יעיל ככל האפשר ולפי הידע שיש לנו כעת בגרסת האלפא זמן הריצה אמור להיות בטווח של שניות עד מספר דקות.

ברגע שהתוכנה תסיים למצוא את כל הפריימרים המתאימים, יטען דף חדש שיציג לו רשימת פריימרים המסודרים בסדר יורד לפי הציון שקיבלו, לכל פריימר שיופק יהיה דוח קצר שיסביר למשתמש באילו טווחים של האילוצים יוצר הפריימר הנ"ל.

לאחר יצירת הפריימר נבדוק האם קיימת אפשרות לבצע שמירה של הפריימר בDB כדי למנוע עבודה מיותרת של התוכנה ובזבוז זמן על חישובים חוזרים לגנים שכבר יוצר עבורם פריימר.

תיאור הכלים המשמשים לפתרון

שפת התכנות בה נשתמש:

- אנו נשתמש בשפה JAVA SCRIPT לכתיבת התוכנה ואתר ה WEB.
- נכון לגרסת האלפא התוכנה נכתבת כתוכנה לוקאלית בשפת PYTHON 2.7, בשלב מאוחר יותר ננסה להפוך אותה לתוכנת WEB.
- התוכנה נכתבת כרגע בתוכנת PYCHARM ומנוהלת ע"י מאגר קוד GITHUB.

סקירת עבודות דומות בספרות והשוואה

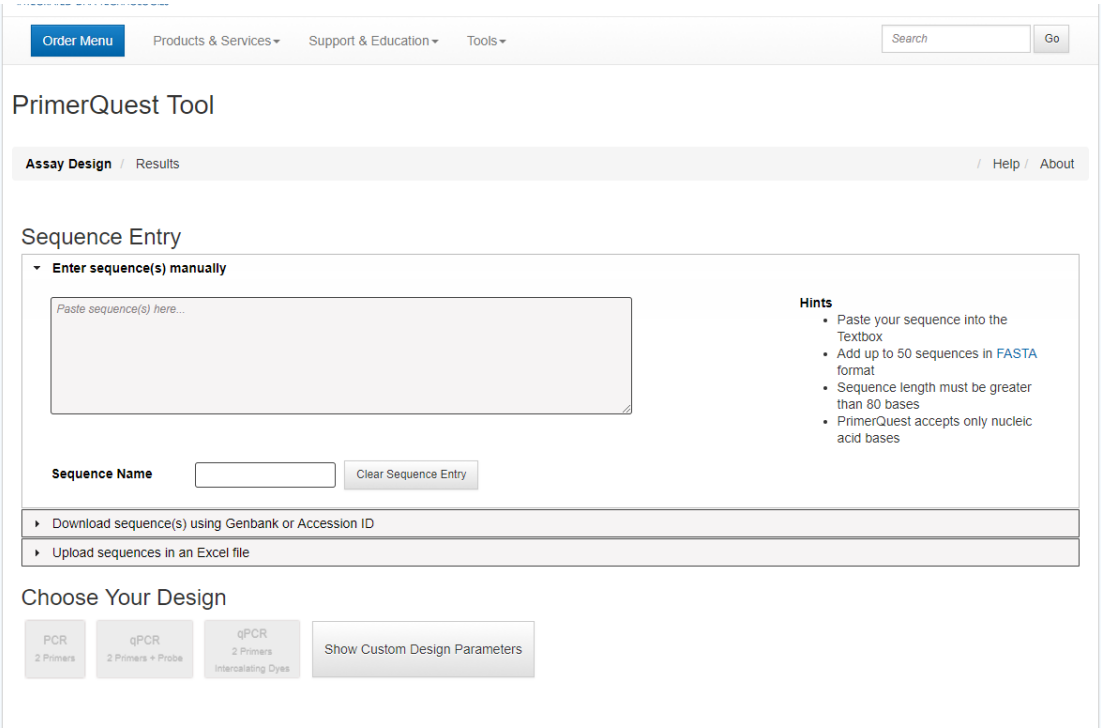
בפרק זה נבחן את המוצרים הדומים הקיימים כיום בשוק:

1. IDT - PrimerQuest Tool :
אתר של חברת סינזטה המספק שירותים שונים הקשורים ל-DNA.
האתר מציע ממשק לבניית פריימר לפי קבלת רצף או שם של גן ובניית פריימר בהתאם לטווחי האילוצים שקובע המשתמש בצורה ידנית.
האתר מציג מספר של פריימרים (אפשר לקבוע את הכמות) ומציג פרטים לגבי הפריימרים:
הפריימר עצמו (FORWARD,REVERSE) , אורך הפריימר , נקודת התחלה וסיום , אחוז קשרי GC , TMו .

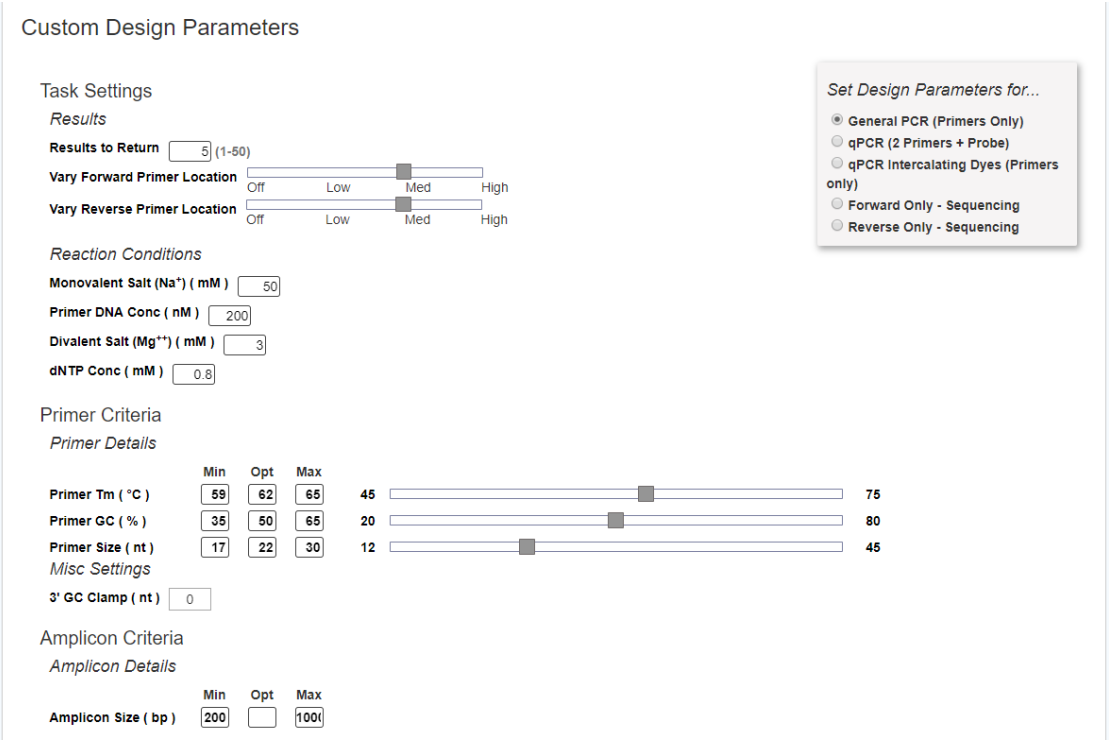
קישור : https://eu.idtdna.com/Primerquest/Home/Index

מסקנות: האתר מציע כביכול את השירות אותו אנו רוצים לתת , אך דורש מהמשתמש לבצע שינויים באילוצים בצורה ידנית ובכך יכול להיווצר מצב שלא ימצא פריימר מהסיבה שהוא אינו נמצא בטווחים שהכניס המשתמש למרות שקיים פריימר מתאים.
בנוסף, אין התייחסות למוטציות הידועות של הגן, יכול להיות שנקבל פריימר שמכיל מוטציה ולכן לא יתאים לכלל האוכלוסייה.

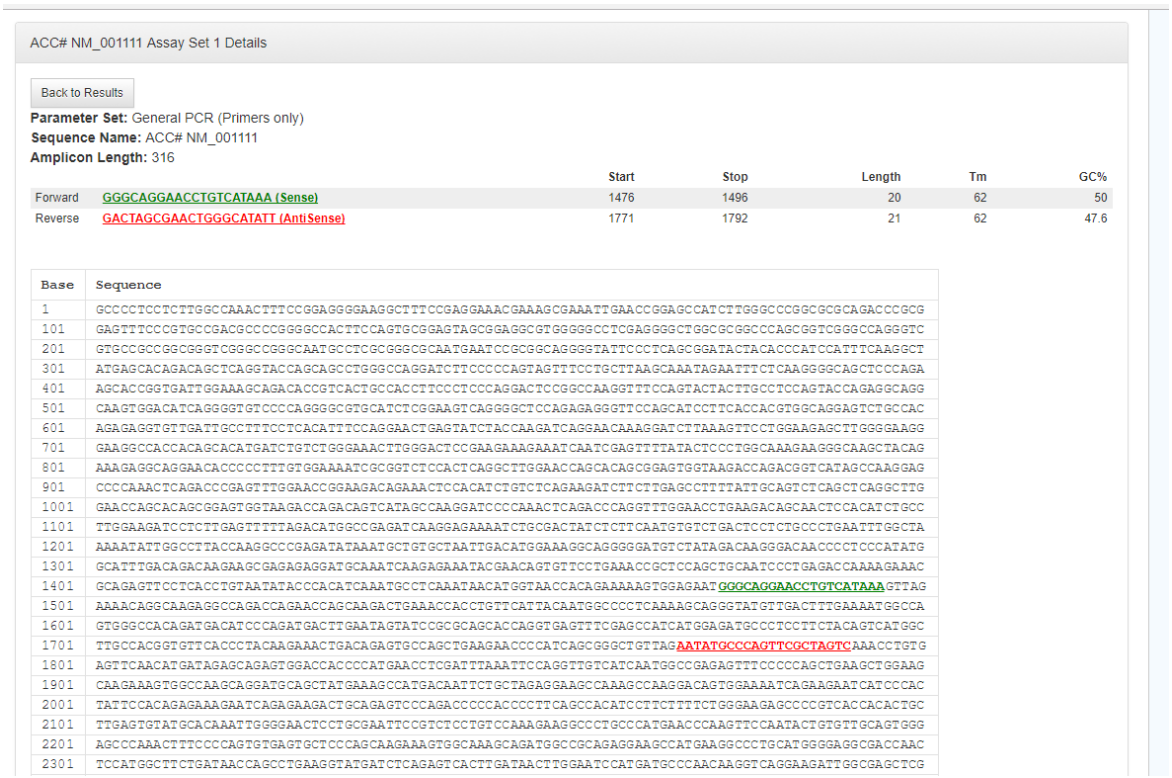
השראה : ממשק משתמש נוח בעיקר בהצגת התוצאות, קבלת תוצאות מהירה (שניות).



מסך חיפוש הגן לפי שם או רצף



מסך שינוי הפרמטרים לאילוצים



ACC# NM_001111 Assay Set 1 Details

Back to Results

Parameter Set: General PCR (Primers only)
Sequence Name: ACC# NM_001111
Amplicon Length: 316

		Start	Stop	Length	Tm	GC%
Forward	GGGCAGGAACCTGTCATAAA (Sense)	1476	1496	20	62	50
Reverse	GACTAGCGAACTGGGCATATT (AntiSense)	1771	1792	21	62	47.6

Base	Sequence
1	GCCCCTCCTCTTGGCCAAACTTTCCGGAGGGGAAGGCTTTCCGAGGAAACGAAAGCGAAATTGAACCGGAGCCATCTTGGGCCCGGCGCGCAGACCCGCG
101	GAGTTTCCCGTGCCGACGCCCCGGGGCCACTTCCAGTGCGGAGTAGCGGAGGCGTGGGGGCCTCGAGGGGCTGGCGCGGCCCAGCGGTCGGGCCAGGGTC
201	GTGCCGCCGGCGGGTCGGGCCGGGCAATGCCTCGCGGGCGCAATGAATCCGCGGCAGGGGTATTCCCTCAGCGGATACTACACCCATCCATTTCAAGGCT
301	ATGAGCACAGACAGCTCAGGTACCAGCAGCCTGGGCCAGGATCTTCCCCCAGTAGTTTCCTGCTTAAGCAAATAGAATTTCTCAAGGGGCAGCTCCCAGA
401	AGCACCGGTGATTGGAAAGCAGACACCGTCACTGCCACCTTCCCTCCCAGGACTCCGGCCAAGGTTTCCAGTACTACTTGCCTCCAGTACCAGAGGCAGG
501	CAAGTGGACATCAGGGGTGTCCCCAGGGGCGTGCATCTCGGAAGTCAGGGGCTCCAGAGAGGGTTCCAGCATCCTTCACCACGTGGCAGGAGTCTGCCAC
601	AGAGAGGTGTTGATTGCCTTTCCTCACATTTCCAGGAACTGAGTATCTACCAAGATCAGGAACAAAGGATCTTAAAGTTCCTGGAAGAGCTTGGGGAAGG
701	GAAGGCCACCACAGCACATGATCTGTCTGGGAAACTTGGGACTCCGAAGAAAGAAATCAATCGAGTTTTATACTCCCTGGCAAAGAAGGGCAAGCTACAG
801	AAAGAGGCAGGAACACCCCCTTTGTGGAAAATCGCGGTCTCCACTCAGGCTTGGAACCAGCACAGCGGAGTGGTAAGACCAGACGGTCATAGCCAAGGAG
901	CCCCAAACTCAGACCCGAGTTTGGAACCGGAAGACAGAAACTCCACATCTGTCTCAGAAGATCTTCTTGAGCCTTTTATTGCAGTCTCAGCTCAGGCTTG
1001	GAACCAGCACAGCGGAGTGGTAAGACCAGACAGTCATAGCCAAGGATCCCCAAACTCAGACCCAGGTTTGGAACCTGAAGACAGCAACTCCACATCTGCC
1101	TTGGAAGATCCTCTTGAGTTTTTAGACATGGCCGAGATCAAGGAGAAAATCTGCGACTATCTCTTCAATGTGTCTGACTCCTCTGCCCTGAATTTGGCTA
1201	AAAATATTGGCCTTACCAAGGCCCGAGATATAAATGCTGTGCTAATTGACATGGAAAGGCAGGGGGATGTCTATAGACAAGGGACAACCCCTCCCATATG
1301	GCATTTGACAGACAAGAAGCGAGAGAGGATGCAAATCAAGAGAAATACGAACAGTGTTCCTGAAACCGCTCCAGCTGCAATCCCTGAGACCAAAAGAAAC
1401	GCAGAGTTCCTCACCTGTAATATACCCACATCAAATGCCTCAAATAACATGGTAACCACAGAAAAAGTGGAGAATGGGCAGGAACCTGTCATAAAGTTAG
1501	AAAACAGGCAAGAGGCCAGACCAGAACCAGCAAGACTGAAACCACCTGTTCATTACAATGGCCCCTCAAAAGCAGGGTATGTTGACTTTGAAAATGGCCA
1601	GTGGGCCACAGATGACATCCCAGATGACTTGAATAGTATCCGCGCAGCACCAGGTGAGTTTCGAGCCATCATGGAGATGCCCTCCTTCTACAGTCATGGC
1701	TTGCCACGGTGTTCACCCTACAAGAAACTGACAGAGTGCCAGCTGAAGAACCCCATCAGCGGGCTGTTAGAATATGCCCAGTTCGCTAGTCAAACCTGTG
1801	AGTTCAACATGATAGAGCAGAGTGGACCACCCCATGAACCTCGATTTAAATTCCAGGTTGTCATCAATGGCCGAGAGTTTCCCCCAGCTGAAGCTGGAAG
1901	CAAGAAAGTGGCCAAGCAGGATGCAGCTATGAAAGCCATGACAATTCTGCTAGAGGAAGCCAAAGCCAAGGACAGTGGAAAATCAGAAGAATCATCCCAC
2001	TATTCCACAGAGAAAGAATCAGAGAAGACTGCAGAGTCCCAGACCCCCACCCCTTCAGCCACATCCTTCTTTCTGGGAAGAGCCCCGTCACCACACTGC
2101	TTGAGTGTATGCACAAATTGGGGAACTCCTGCGAATTCCGTCTCCTGTCCAAAGAAGGCCCTGCCCATGAACCCAAGTTCCAATACTGTGTTGCAGTGGG
2201	AGCCCAAACTTTCCCCAGTGTGAGTGCTCCCAGCAAGAAAGTGGCAAAGCAGATGGCCGCAGAGGAAGCCATGAAGGCCCTGCATGGGGAGGCGACCAAC
2301	TCCATGGCTTCTGATAACCAGCCTGAAGGTATGATCTCAGAGTCACTTGATAACTTGGAATCCATGATGCCCAACAAGGTCAGGAAGATTGGCGAGCTCG

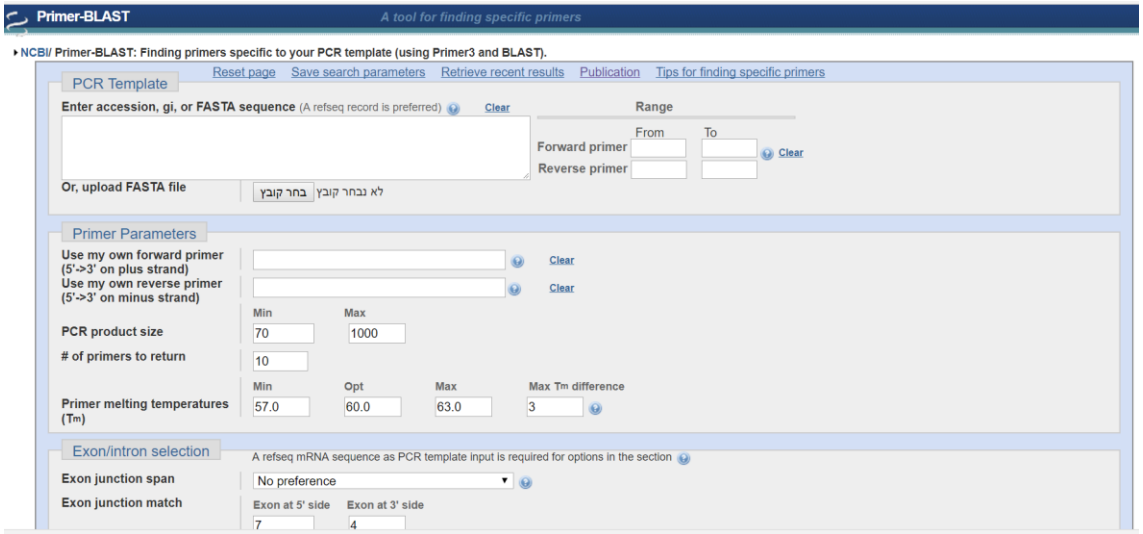
מסך תצוגת הפריימר שנמצא

2. NCBI PRIMER-BLAST

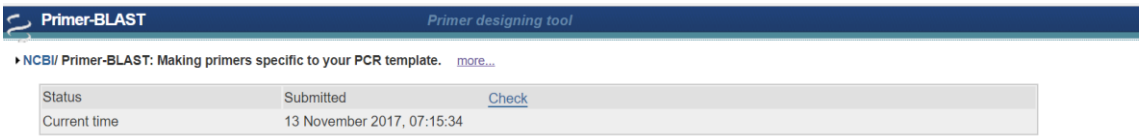
אתר NCBI – National Center for Biotechnology Information הוא אתר הקשור למכונים הלאומיים לבריאות של ארה"ב, הוא מספק מידע הכולל מאמרים בנושאים מחקריים ביורפואיים, ומחזיק מאגרי מידע נוספים הרלוונטיים למחקר בנושא הביוטכנולוגיה, וביניהם מידע על DNA.
באתר קיים כלי שנקרא PRIMER BLAST – בכלי זה ניתן לבנות פריימר ע"י הכנסת gene sequence והכנסת הפרמטרים הרצויים הקשורים לאילוצים כמו טמפ', אורך, אחוז קשרי ה GC, האם הפריימר חייב להיות מורכב מחיבורים של exon-ים או לא וכו'.
לבסוף האתר מציג את הפריימר.

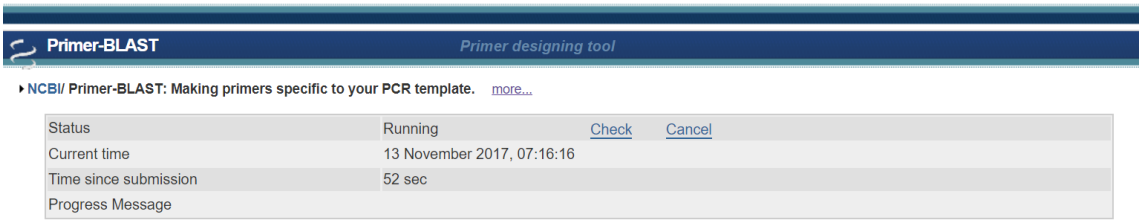
קישור: https://www.ncbi.nlm.nih.gov/tools/primer-blast/

השראה: באתר הזה ראינו שלאחר שהרצנו את החיפוש הרצוי, הופיע מסך שהציג את הזמן בו התחלנו את החיפוש והציג את סטטוס הפנייה ואיפה עומד התהליך.



מסך חיפוש הפריימר לפי שם או רצף ושינוי הפרמטרים





מסך עדכון תהליך יצירת הפריימר

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TCCGAGGAAACGAAAGCGAA	Plus	20	40	59	59.97	50.00	4.00	0.00
Reverse primer	GGAGCTGCCCCTTGAGAAAT	Minus	20	396	377	60.03	55.00	5.00	2.00
Product length	357								

Products on intended target

>NM_001111.4 Homo sapiens adenosine deaminase, RNA specific (ADAR), transcript variant 1, mRNA

```
product length = 357
Forward primer  1   TCCGAGGAAACGAAAGCGAA  20
Template        40  ....................  59

Reverse primer  1    GGAGCTGCCCCTTGAGAAAT  20
Template        396  ....................  377
```

Products on potentially unintended templates

>NM_015841.3 Homo sapiens adenosine deaminase, RNA specific (ADAR), transcript variant 3, mRNA

```
product length = 357
Forward primer  1   TCCGAGGAAACGAAAGCGAA  20
Template        40  ....................  59

Reverse primer  1    GGAGCTGCCCCTTGAGAAAT  20
Template        396  ....................  377
```

>NM_015840.3 Homo sapiens adenosine deaminase, RNA specific (ADAR), transcript variant 2, mRNA

```
product length = 357
Forward primer  1   TCCGAGGAAACGAAAGCGAA  20
Template        40  ....................  59

Reverse primer  1    GGAGCTGCCCCTTGAGAAAT  20
Template        396  ....................  377
```

>XM_006711111.3 PREDICTED: Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant X5, mRNA

```
product length = 357
Forward primer  1    TCCGAGGAAACGAAAGCGAA  20
Template        104  ....................  123

Reverse primer  1    GGAGCTGCCCCTTGAGAAAT  20
Template        460  ....................  441
```

ישנן עוד מספר אפליקציות רשת העובדות בצורה זהה :

- http://biotools.umassmed.edu/bioapps/primer3_www.cgi
- http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi
- http://bisearch.enzim.hu/?m=search

כולם עובדים בצורה זהה של הכנסת רצף/שם הגן , וקביעת פרמטרים.
אף אחד מהכלים לא מציע מציאת פריימר ללא התערבות בפרמטרים ובהתחשבות במוטציות והתחשבות בפריימר ש"ישב" על נקודת חיבור של 2 אקסונים (Junctions).

תכנית בדיקות

גורם נבדק	תוצאת בדיקה רצויה	תהליך ביצוע הבדיקה	תיקון במידה ולא התקבלה תוצאה רצויה
בדיקת הפונקציה Get_varient(species,symbol) הפונקציה מקבלת סוג(אדם, עכבר וכו')את שם הגן ומחזירה אובייקט json המכיל את כל הפרטים על הגן.	אובייקט json המכיל את המידע על הגן שנמצא באתר ensembl	לאחר הרצת הפונקציה ייצאנו את הפלט לקובץ וביצענו השוואה שאכן קיימים כל הפרטים הנחוצים על הגן ושהם אכן תואמים לגן שנמצא באתר ensembl ממנו	
בדיקת חיתוך ה transcript לtranscript הנחוץ, לפי הנחיות הביולוג אנו צריכים לקבל את ה transcript שמתחיל בשמו של הגן ובהמשך מופיע 201- למשל עבור הגן BRAF אנו רוצים לקבל מתוך אוביקט הjson את הtranscript: BRAF-201.	אובייקט json של ה transcript 201- הרצוי, המכיל את המידע שנמצא באתר Ensembl	לאחר קבלת אובייקט הjson של ה transcript ביצענו השוואה מול הנתונים באתר Ensembl כדי לראות שהפרטי הtranscript נכונים, קיימים בו הexon-ים הנכונים, אינדקס ההתחלה שלו והסוף שלו תואמים את הנתונים באתר.	
בדיקת הפונקציה junctions(transcript) הפונקציה מקבלת transcript ומחזירה מערך עם נקודות החיבור(junction) בין האקסונים	מערך שמכיל נקודות שהמעבר שכאשר עוברים בין אות שיושבת בנקודה המסוימת הזאת בtranscript לאות הבאה בעצם עברנו בין אקסון אחד לאקסון אחר, כמו כן שלא חרגנו מהtranscript.	לאחר הרצת הפונקציה בדקנו בעזרת האתר שאכן בנקודות החיבור של האקסונים שמצאנו מתבצע המעבר בין אקסון לאקסון, הבדיקה התבצעה על 2 גנים שונים, BRAF ו BRCA2	לקח לנו מספר פעמים לרשום את הפונקציה משום שבהתחלה היו לנו בעיות של אי התאמה בין מה שמצאנו לבין מה שהיה קיים באתר ז"א שנמצאו נקודות שלא היו בחיבור בין האקסונים, ביצענו שינויים בחישוב עד שהגענו לתוצאה שעובדת בצורה תקינה.

בדיקת הפונקציה get_cdna(transcript id) ההפונקציה מקבלת את ה id של ה transcript ומחזירה את ה cdna (רצף אותיות) של ה transcript.	אנו מצפים לקבל בחזרה cdna (רצף אותיות) זהה לחלוטין לרצף הקיים באתר ensemble שמתאים ל transcript המסוים של הגן המסוים שבחרנו.	לאחר הרצת הפונקציה ביצענו הדפסה של ה cdna ובעזרת פונקציית השוואת מחרוזות ביצענו השוואה בין ה cdnaשקיבלנו לבין זה שמופיע באתר.	גילינו שהפונקציה שלנו החזירה את ה cdna שכולל גם ירידת שורה לכן ביצענו לאחר מכן על הפלט החלפה של כל ירידת שורה לתו ריק כדי ליצור רצף ללא רווחים וירידת שורות
בדיקת הפונקציה Get_optional_primers(cdna,juncArr,id_count) הפונקציה מקבלת את ה Cdna ואת המערך שמכיל את כל ה junctions וcounter כדי להעניק ID ייחודי לכל פריימר ומחזירה רשימת פריימרים באורכים שנקבעו בקובץ הקונפיגורציה שיושבים על חיבורים של אקסונים(junctions)	מערך שמכיל פריימרים באורכים שנקבעו בקובץ הקקונפיגורציה ויושבים על חיבורים של אקסונים.	לאחר הרצת הפונקציה בדקנו מספר פריימרים האם הם באמת באורך הנכון והאם הום באמת יושבים על חיבורים של אקסונים ביחס הנכון(40%-60%).	
בדיקת הפונקציה primer_tests(optional_primers) הפונקציה מקבלת מערך של פריימרים אופציונליים ומחזירה את הפריימרים שעומדים בטווחי הפרמטרים שניתנו בקובץ הקונפיגורציה ובדיקות רצפים שונות של אותיות שאסור שיהיו בפריימר	מערך שמכיל את הפריימרים שעברו את הסינון של האילוצים: GC,TM, ובדיקות רצפים שונות של אותיות שאסור שיהיו בפריימר.	לאחר ההרצה ביצענו הדפסה של הפריימרים ובדקנו מדגמית מספר פריימרים שאכן הם עומדים בתנאים שנבדקו.	
בדיקת הפונקציה get_reverse_primers(cDna,forward_primers,id,count) הפונקציה מקבלת את ה CDNA ואת רשימת ה forward_primers ואת ה id_count כדי להעניק ID ייחודי לכל פריימר, הפונקציה מחזירה רשימת reverse primers כאשר לכל פריימר ברשימה מותאם זוג forward כל reverse נמצא על אמפליקון בטווח האורך שנקבע בקובץ הקונפיגורציה	מערך שמכיל reverse primers שכל אחד מהם מתאים לforward פריימר ויושב על אמפילקון בטווח האורך שנקבע בקובץ הקונפיגורציה	לאחר הוצאת reverse primers ביצענו הדפסה של זוגות הפריימרים ובדקנו שאכן שני הפריימרים יושבים על אותו אמפליקון.	
בדיקת הפונקציה Get_optional_sets(forward_primers,reverse_primers) הפונקציה מקבלת את רשימת ה forward וה reverse Primers ויוצאת מהם את הזוגות לפי התאמות בין ה IDי-ם שהוגדרו.	מערך שמכיל בכל תא אובייקט של זוג פריימרים (forward,reverse) שתואמים אחד לשני לפי ה ID שנשמר ב reverse פריימר	לאחר הוצאת Primer sets ביצענו הדפסה של כל אובייקט זוג פריימרים ובדקנו שאכן ה ID שקיים ל forward תואם ל id שקיים ב pair id אצל ה reverse.	
בדיקת הפונקציה Sets_tests(primers_optional_sets) הפונקציה מקבלת זוגות אופציונליים ובודקת את הפרש הטמפרטורה ביניהם ומחזירה מערך שמכיל את כל הזוגות שהפרש הטמפרטורה שלהם נמצא בטווח המוגדר בקובץ הקונפיגורציה.	מערך המכיל בכל תא אבוייקט של זוג פריימרים שהפרש התמפרטורה שלהם נמצא בטווח הקיים בקובץ הקונפיגורציה	לאחר הרצת הפונקציה ביצענו הדפסה של כל אובייקט זוג פריימרים, בהדפסה הדפסנו גם את מאפיין הפרש הטמפרטורה ובדקנו שאכן ההפרש בטווח המוגדר.	

בדיקת הפונקציה Export_to_file(primers_sets,species,symbol) הפונקציה מקבלת מערך עם זוגות הפריימרים את שם הזן ושם הגן ומוציאה לקובץ את 100 זוגות הפריימרים עם הציון הגבוהה ביותר עם הפרטים שעליהם	קובץ עם 100 זוגות הפריימרים עם הציון הגבוהה ביותר עם הפרטים שעליהם, הקובץ מכיל בכותרת את שם הזן ושם הגן.	עברנו על הקובץ ובדקנו שהפריימרים שמופיעים במערך אכן מופיעים בקובץ וממוינים ע"פ הציון.	

סיכום/מסקנות - גרסת אלפא:

מצב נוכחי:

גרסת האלפא שאנו מגישים ב1.2.2018 מכילה את הפיצ'רים הבאים:

- קבלת שם גן מהמשתמש ומשיכת המידע הנחוץ עליו (הרצף ומיקומי האקסונים) מהDATABASE של אתר ENSEMBEL.
- פעולת האלגוריתם למציאת פריימרים מתאימים שכוללת בשלב זה את רוב המבחנים והענקת ציון לכל זוג פריימרים שנמצא (פירוט בהמשך).
- יצירת קובץ OUTPUT המכיל רשימה של 100 זוגות הפריימרים עם הציון הגבוהה ביותר מסודרים בסדר יורד.
- התוכנה הנוכחית כתובה בPYTHON ועובדת לוקאלית בלבד, יש צורך בהתקנת PYTHON במחשב ובהרצה מהCMD.

פירוט על הקוד:
מחלקות:

1. Primer.py – המחלקה מייצגת אובייקט מסוג PRIMER , כל אובייקט מכיל את המאפיינים השונים על הפריימר כגון : אורך , טמפ' פרימה , אחוז GC ועוד...
2. Primer_set.py – המחלקה מייצגת זוגות פריימרים ומכילה את המידע הרלוונטי לכל זוג, הזוגות שנוצרים הינם זוגות שעברו את המבחנים ולכן כל זוג גם יקבל ציון.
3. EnsembelRestClient.py – מחלקה שמבצעת את הפניות לשרת API של אתר ENSEMBEL ומחזירה את המידע המבוקש על הגן.
4. Primer_builder.py – המחלקה הראשית שמכילה גם את הMAIN , מרכיבה את זוגות הפריימרים בעזרת פונקציות של מבחנים ושימוש במחלקות הנ"ל.

תהליך הריצה:

1. MAIN – המשתמש מכניס את שם הגן ושם הזן לדוגמא : BRAF ו-human .
2. בקשה מה DATABASE וקבלת אובייגקט JSON המכיל את כל הפרטים על הגן.
3. פירוק אובייקט הJSON ומשיכת הרצף של הטרנסקריפט הראשי של הגן ושל מיקומי האקסונים.
4. הכנסת רצף הDNA לרשימה (מערך) ויצירת רשימה של מיקומי חיבורי האקסונים.
5. יצירת כל הפריימרים FORWARD בכל האורכים האפשריים שנמצאים על הJUNKTIONS (נקודות חיבור האקסונים).
6. שליחת הפריימרים שנמצאו למבחנים ופסילת הפריימרים שלא עומדים בתנאי הקיצון. מבחנים סינטקטיים (רצפים אסורים) , מבחני טמפ' וכד'.
7. יצירת כל פריימרי הREVERSE בכל אורך אפשרי שנמצאים במרחק תקין מכל הFORWARD שנמצאו.
8. שליחת פריימרי הREVERSE למבחנים(כמו בסעיף 6).
9. צימוד הזוגות שעברנו את המבחנים לזוגות פריימרים מתאימים.
10. נתינת ציונים לזוגות הפריימרים ע"י נוסחאת ציון (לא סופית כרגע).
11. הדפסת 100 הזוגות בעלי הציון הגבוה ביותר לקובץ TXT בתוך תיקייה נפרדת שנקרת output.

דוגמת הרצה של תוכנה בגרסה הנוכחית:

```
C:\Users\Amir\Desktop\Amir\Final project\DNA-PRIMER-BUILDER\PythonServer>python Primer_Builder.py
Enter specie:Human
Enter symbol:BRAF
Proceed to the output folder to view the results.
```

פלט התוכנה:

```
PRIMERS SETS of BRAF(Human Gene):
--------------------------------------------------------------------------------
FORWARD :
Sequence: AGTGCTGTGCTGTTTACAG length: 19 Tm: 48 GC: 47 start index: 654
REVERSE:
Sequence: ACAACACACAACTTTGTACG length: 20 Tm: 47 GC: 40 start index: 773
tm difference between F & R: 1
Amplicon length: 119
Score: 40
--------------------------------------------------------------------------------
FORWARD :
Sequence: AGTGCTGTGCTGTTTACAG length: 19 Tm: 48 GC: 47 start index: 654
REVERSE:
Sequence: CAACACACAACTTTGTACGA length: 20 Tm: 47 GC: 40 start index: 774
tm difference between F & R: 1
Amplicon length: 120
Score: 40
--------------------------------------------------------------------------------
FORWARD :
Sequence: ATCATTTGGAACAGTCTAC length: 19 Tm: 44 GC: 36 start index: 1474
REVERSE:
Sequence: AAAACACGACATGTGAATAT length: 20 Tm: 43 GC: 30 start index: 1595
tm difference between F & R: 1
Amplicon length: 121
Score: 40
--------------------------------------------------------------------------------
FORWARD :
Sequence: ATCATTTGGAACAGTCTAC length: 19 Tm: 44 GC: 36 start index: 1474
REVERSE:
Sequence: AAACACGACATGTGAATATC length: 20 Tm: 45 GC: 35 start index: 1596
tm difference between F & R: 1
Amplicon length: 122
Score: 40
```

מסקנות :

1. המחסום הביולוגי – בשלב זה המחסום הגדול ביותר הוא החוסר בידע ואפלוליות בתחום הביולוגי, אנו נתקלים בקשיים וחוסר הבנה רבים בתחום זה שגורמים לעיכובים ותקיעות בהמשך העבודה על הפרויקט.
נוכחנו שהתחום עליו אנו עובדים הוא תחום קצת אפור ולא מדע מדויק, טווחי הפרמטרים משתנים מתוכנה לתוכנה ותלויים במשתנים רבים.
ננסה בהמשך הפרויקט להתמקד בדרישות ספציפיות יותר שנקבל מאיש הקשר שלנו מתחום הביולוגיה.

2. בנק פריימרים – לאחר חיפושים רבים התברר לנו שלא קיים בנק פריימרים מהסוג שאנחנו מחפשים שמכיל פריימרים שנוסו ועבדו.
ננסה למצוא מאגר כלשהו של פריימרים או שניצור מאגר של פריימרים שנבקש מהביולוג מולו אנו עובדים בכדי לבצע
reverse engineering ולנתח את המאגר ולהבין ממנו יותר על הפרמטרים והטווחים שלהם.

3. סביבת עבודה – בהגדרה הראשונית הפרויקט היה אמור לעלות כאפליקציית WEB שתכתב בJS , בהמשך עברנו לעבודה בPYTHON בשל הנוחות עבודה עם סטרינגים.
אנו עדיין שואפים שהתוכנה תעלה לרשת לכן ננסה לברר אופציה של המרת הקוד בדרך כלשהי לקוד WEB.

בשלב הבא של הפרויקט נרצה לבצע את הדברים הבאים:

- בדיקת הימצאות הפריימר ב DNA – נרצה לקבל כמות הימצאות ממש קטנה כי אנו רוצים שהפריימר יתאים רק לגן מסוים ולא לעוד קטעים ב DNA.
- הכנת נוסחה מדויקת למדידת הטמפרטורה של הפריימר.
- **נקבל בנק של פריימרים מהביולוג, בעזרת בנק הפריימרים נוכל לבצע בדיקות תקינות ולהוציא נתונים בעזרת הבנק.**
- נרצה להוסיף בדיקה של מבנים שניוניים – הכוונה שהפריימר לא יהיה פלינדרום או פלינדרום חלקי דבר שעלול לגרום להידבקות של 2 קצוות הפריימר (לדוגמא AGTCT הוא מבנה לא טוב כי הרצפים בהתחלה ובסוף עלולים להידבק ולקפל את הפריימר)

נספחים

ספרות, תרשימים נוספים, תכנון הפרויקט, טבלת ניהול סיכונים, טבלת דרישות (URD),

רשימת ספרות \ ביבליוגרפיה

1. Primer express, User guide –
הוראות שימוש לתוכנה לעיצוב פריימרים לPCR , ממנה לקחנו את התנאים לאילוצים.
http://www.bu.edu/picf/files/2010/11/Primer-express-30.pdf

2. A systematic guideline for developing the best real-time PCR primers –
https://www.google.co.il/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0ahUKEwjov5C9pd_YAhUJLZoKHcwXAGMQFghaMAM&url=https%3A%2F%2Fwww.qiagen.com%2Fus%2Fresources%2Fdownload.aspx%3Fid%3Dd6191d0e-701b-4eb1-bafa-d7ab7677875f%26lang%3Den&usg=AOvVaw2HmJpO-elrjx5vj9C4tTly

תכנון הפרויקט

פגישה ראשונה עם דר' חסין לקבלת הפרוייקט	.24.10
הגשת טופס הצעה לאחר אישור המנחה	19.11
סיום מחקר ספרותי	1.12
קוד ראשוני לאלגוריתם , תחילת בדיקות הרצה	20.12
התחלת בניית ממשק משתמש	1.1
הצגת אב טיפוס (אלגוריתם חלקי)	21.1
סיום כתיבת האלגוריתם , תחילת בדיקות	1.3
סיום בדיקות עצמאיות , הפצת גרסת בטא	1.4
קבלת משובים	1.5
סיום תיקונים	20.5

22.6	הגשת הפרוייקט

טבלת סיכונים

1 - נמוך
3 - גבוהה

#	הסיכון	חומרה	סבירות	מענה אפשרי
1	תוכנות מתחרות – תוכנות המציעות את אותו השירות בדיוק.	1	2	ביצוע סקר שוק מקיף לשלילת מוצרים זהים
2	חוסר ידע בתחום הביולוגי – לפרוייקט הקשרים לתחום הביולוגיה ואנו חסרי ידע מקדים בנושא	1	3	נתמקד בהיבטים התוכניתיים ולא הביולוגים של הבעיה ופתרונה. למשל הסתכלות על DNA כסטרינג.
3	עבודה עם מסדי נתונים קיימים – עלולות להית קריסות במסד נתונים שלא תלוי בנו	3	1	ננסה להסתמך על יותר ממקור אחד.
4	ביצועים - לתוכנה יקח זמן רב מידי להפיק את המידע הרצוי	3	2	ננסה לכתוב את האלגוריתם בדיקה היעיל ביותר. כמו כן נאמר לנו שהאלגוריתם אמור לרוץ בטווח של שעות עד ימים – טווח ארוך מספיק.
5	אילוצים רבים מידי – לא נצליח למצוא פריימר התואם לכל האילוצים	2	1	נפנה לחוקר לצורך חידודים באילוצים, כמו כן נוודה שהאלגוריתם מוצא פריימר על גן שידוע שקיים פריימר עבורו
6	עמידה בזמנים – לא נספיק לבנות את התוכנה במועד	2	2	תכנון זמנים יעיל. קביעת ימי עבודה קבועים.
7	מכשולים תכנותיים – חוסר הצלחה בכתיבת האלגוריתם	2	2	ניעזר בגורמים חיצוניים במידת הצורך.
8	תחום הפריימרים הוא תחום "אפור" – בכל מאמר נכתבות המלצות שונות בקשר ליצירת פריימרים, משתמשים בנוסחאות שונות לחישוב ואין ריכוז של פריימרים	2	3	ננסה לשאוב כמה שיותר מידע מהביולוג, ולהיעזר במאמרים, כמו כן ננסה למצוא בנק של פריימרים ולנתח אותו ולעשות reverse engineering של המאפיינים מהבנק.

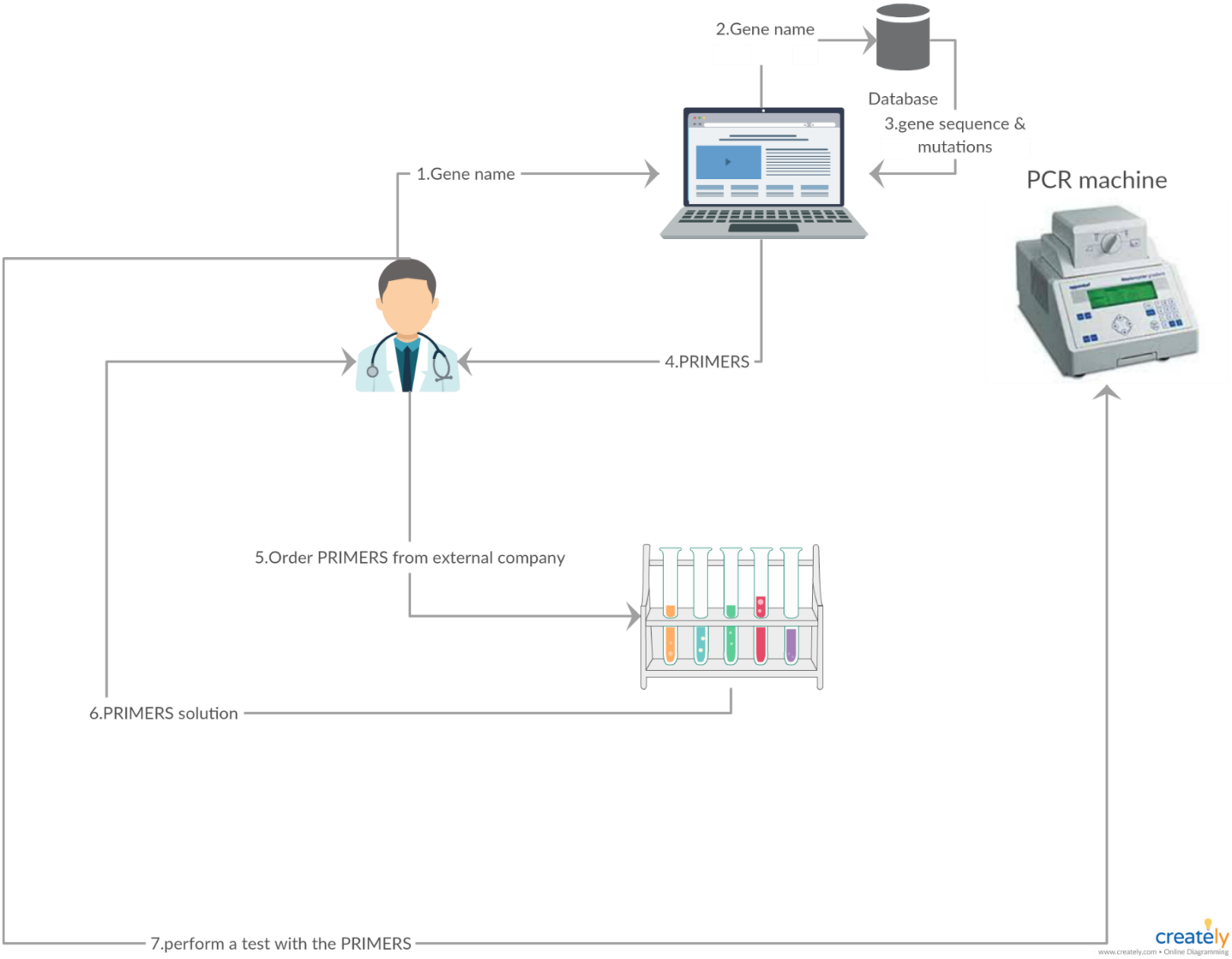
	קיימים שנוצרו ע"י אותם מאפיינים			
9	הכלים התכנותיים שהחלטנו עליהם אינם מתאימים ונאלץ לבצע מיגרציה של הפרויקט	1	2	ננסה להחליט בתחילת הפרויקט ושפות דינאמיות, וננסה לאבחן איזה שפה מתאימה לפרויקט הנ"ל. (במהלך גרסת האלפא יחיסת בהתחלה עברנו javaמ בתוכנת visual studio code ל python בתוכנה pycharm)

טבלת דרישות

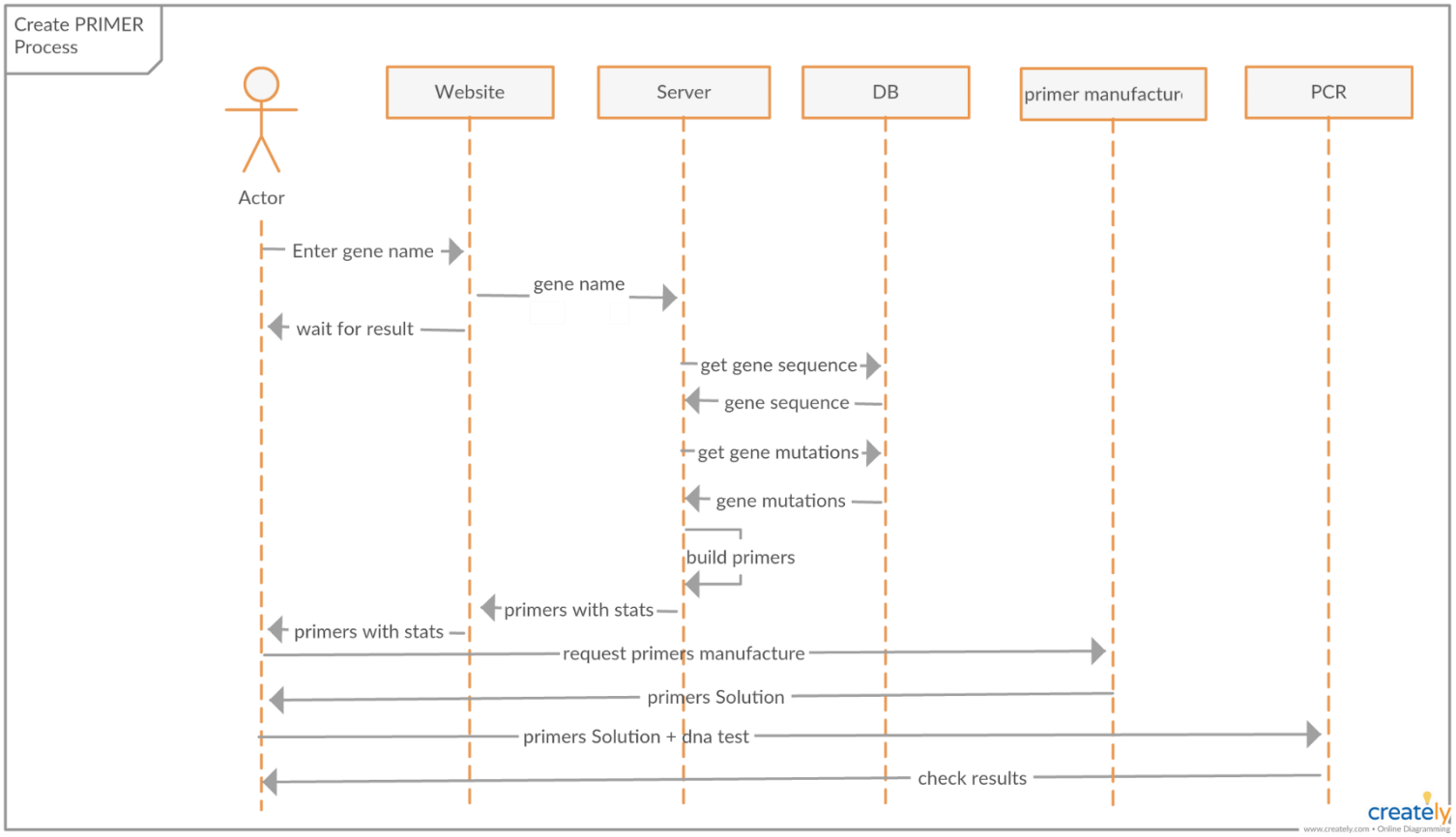
טבלת דרישות (User Requirement Document)

דרישה	תיאור
נכונות	הפריימרים שיתקבלו מהתוכנה יהיו נכונים ואמינים בהתאם לאילוצים
מערכת אוטומטית	מציאת הפריימר תהיה אוטומטית ללא התערבות אדם בפרמטרים
זמנים	התוצאה תתקבל בזמן סביר (כרגע לפי הערכה מס' שעות)
גלובלי	המערכת תהיה זמינה לציבור הרחב (אתר אינטרנט ציבורי)
מס משתמשים	האתר יתמוך במספר בלתי מוגבל של בקשות אך ייתכן שמספר רב של בקשות ירום לעיכובים בתוצאות
קלות שימוש	הממשק יהיה פשוט למשתמש מהתחום, יציג את הנתונים בצורה הברורה ביותר
הצגת התוצאות ברשימה	התוכנה תחזיר פלט למסך\למסמך הכולל את הפריימרים מסודרים בסדר יורד לפי איכות הפריימר ותציג עבור כל פריימר את התכונות שלו.
התוכנה תעבוד מול DB מסוים	התוכנה תעבוד מול מסד הנתונים של אתר ensemble

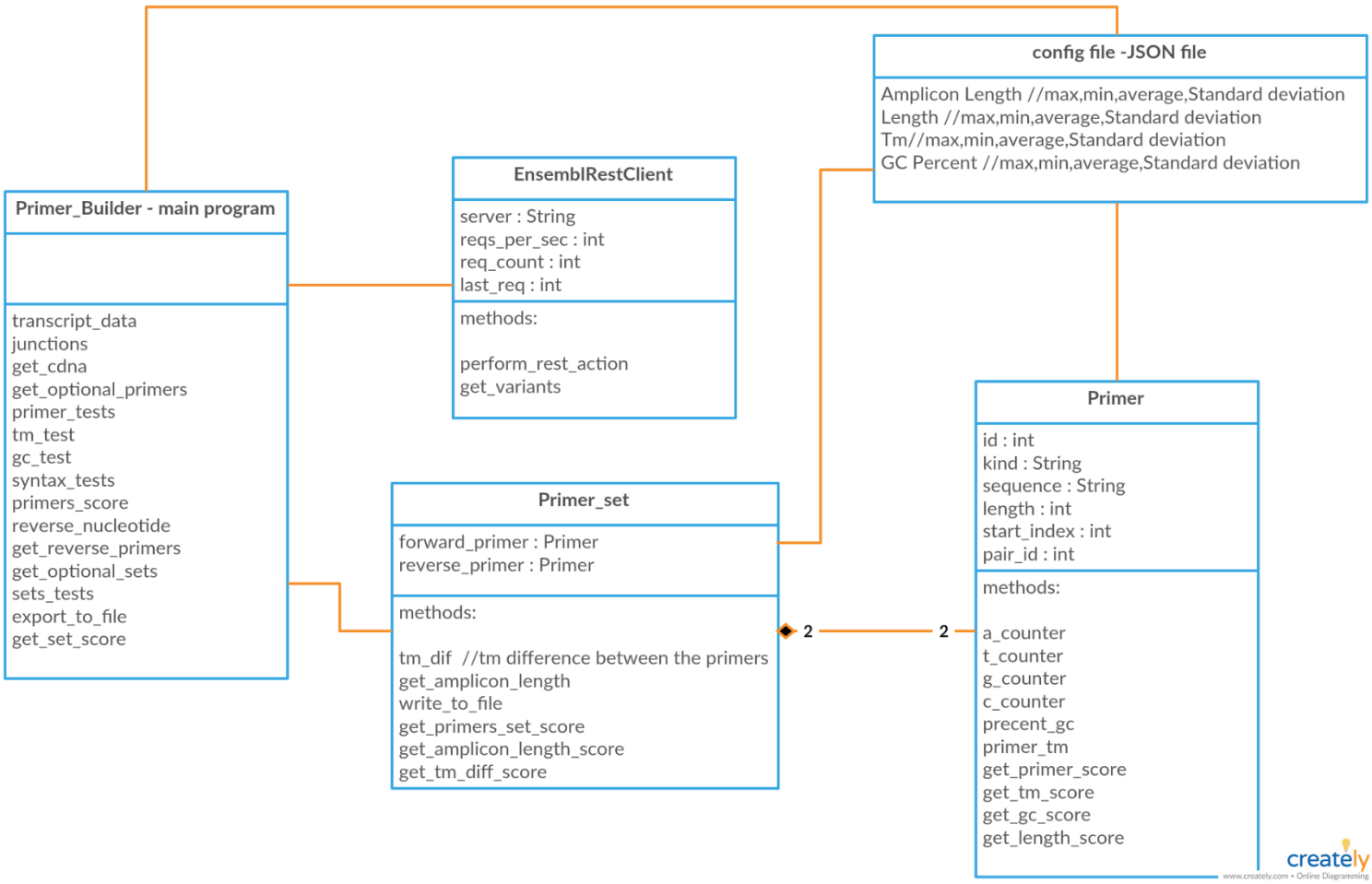
תרשים תפקידים כללי :



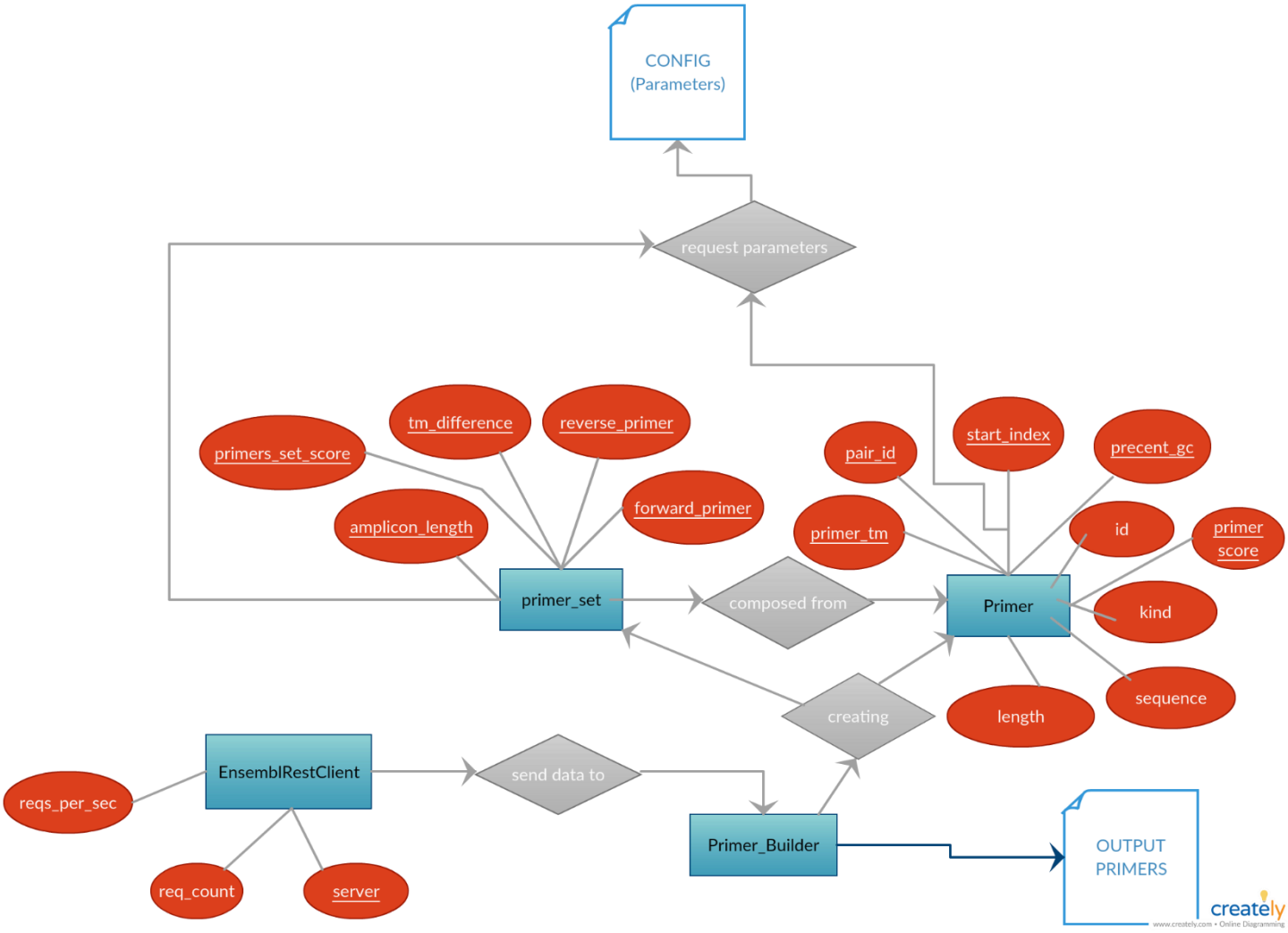
Sequence diagram



Class diagram



ERD


CONFIG
(Parameters)
request parameters
tm_difference
reverse_primer
primers_set_score
amplicon_length
forward_primer
pair_id
start_index
precent_gc
primer_tm
id
primer score
primer_set
composed from
Primer
kind
sequence
length
creating
EnsemblRestClient
send data to
reqs_per_sec
req_count
server
Primer_Builder
OUTPUT
PRIMERS
creately