

# אוניברסיטת בר-אילן

אפיון מולקולארי של שונות סומקלונלית בעצי תמר (*Phoenix dactylifera* L.)

מוגש ע"י ענת רומנובסקי 308998582

עבודה זו מוגשת כחלק מהדרישות לשם קבלת תואר מוסמך  
בפקולטה למדעי החיים ע"ש מינה ואבררד גודמן של אוניברסיטת בר-אילן

עבודה זו נעשתה בהדרכתם של:  
ד"ר יובל כהן, מנהל המחקר החקלאי, מכון וולקני  
פרופ' אורי לביא, מנהל המחקר החקלאי, מכון וולקני  
פרופ' רפאל פרל-טרבס, אוניברסיטת בר אילן.

תודות:

ברצוני להודות לד"ר יובל כוהן, פרופ' אורי לביא ופרופ' רפאל פרל-טרבס על ההזרחה, העזרה הרבה והסבלנות לאורך כל הדרך.

תודה רבה לפרופ' אהרון רזין ושירי על העזרה בעבודה על מתילציה.

תודה רבה לד"ר רון אופיר ושי אוליאל על העזרה בעבודה הביואינפורמטית.

תודה לד"ר שירלי גזית וד"ר Hewan Demissie Degu על שיתוף הפעולה והעזרה בעבודתנו המשותפת במעבדה.

תודה רבה לכל צוות המעבדה ובמיוחד לרעיה קורצ'ינסקי על העזרה, התמיכה והאווירה הנעימה.

תודה מיוחדת לבן זוגי ולדיק ולבני משפחתי על התמיכה, העזרה והסבלנות הרבה במהלך התואר.

## תוכן עניינים

1.....	תוכן עניינים	
3.....	רשימת טבלאות	
4.....	רשימת איורים	
5.....	תקציר	
7.....	1 רקע מדעי	
7.....	1.1 עץ התמר והדרכים לריבויו	
8.....	1.2 שונות סומקלונלית	
8.....	1.3 המנגנון האחראי ליצירת שונות סומקלונלית	
9.....	1.4 שונות סומקלונלית וטיפוסים חריגים בתמר	
10.....	1.5 שיטות מולקולאריות לבחינת שונות סומקלונלית	
12.....	1.6 אפיון מתילציה	
15.....	1.7 אפיון דגם ביטוי גנים	
16.....	2 מטרות המחקר	
17.....	3 שיטות וחומרים	
17.....	3.1 תילציה של DNA	
17.....	3.1.1 החומר הצמחי	
17.....	3.1.2 הפקת DNA גנומי	
18.....	3.1.3 הגברת מקטעים באמצעות PCR	
19.....	3.1.4 בדיקת תוצרי ה-PCR	
19.....	3.1.5 קביעת רצף תוצרי ה-PCR	
20.....	3.1.6 אפיון מתילציה במקטעי DNA בשיטת ה-Bisulfite sequencing	
21.....	3.1.7 שיבוט התוצרים	
21.....	3.1.8 טרנספורמציה לחיידקים	
22.....	3.2 ריצוף בהספק גבוה	
22.....	3.2.1 החומר הצמחי	
23.....	3.2.2 הפקת Total RNA	
24.....	3.2.3 יצירת cDNA	
25.....	3.2.4 ריצוף בשיטת Solexa/Illumina	
25.....	3.2.4.1 Digital Gene Expression (DGE)	
25.....	3.2.4.1.1 עיכול עם <i>NlaIII</i>	
25.....	3.2.4.1.2 עיכול באנזימי רסטריקציה שונים	
26.....	3.2.4.2 mRNA-seq	
26.....	3.2.4.3 תהליך הריצוף	
28.....	3.2.5 נירמול לתוצאות אנליזת ה-DGE	
28.....	3.2.6 5' RACE	

29.....	אנליזת BLAST	3.2.7
30.....	אפיון דגם השעתוק באמצעות Q-RT-PCR	3.2.8
<b>34.....</b>	<b>4 תוצאות</b>	
34.....	אפיון הבדלים בדגם מתילציה ה-DNA	4.1
34.....	ביקורת חיובית לאפיון דגם המתילציה בשיטת Bisulfite sequencing	4.1.1
34.....	אפיון מתילציה של מקטעי DNA בשיטת ה-MS-AFLP	4.1.2
39.....	אפיון דגם המתילציה של הגן DEFICIENS (DEF)	4.1.3
45.....	אפיון דגם הביטוי בעצי תמר ננסיים ובעצי תמר נורמאליים על ידי ריצוף בהספק גבוה	4.2
45.....	אפיון דגם הביטוי באמצעות Digital Gene Expression	4.2.1
45.....	תוצאות ה-Digital Gene Expression	4.2.1.1
46.....	הארכת רצף באמצעות 5' RACE	4.2.1.2
46.....	אנליזת BLAST	4.2.1.3
46.....	אנליזת BLAST כללית	4.2.1.3.1
46.....	אנליזת BLAST לגנום התמר	4.2.1.3.2
49.....	אימות דגם הביטוי של ה-DGE באמצעות Q-RT-PCR	4.2.1.4
51.....	אפיון דגם הביטוי באמצעות mRNA-seq	4.2.2
51.....	תוצאות ה-mRNA-seq	4.2.2.1
51.....	סינון התוצאות	4.2.2.2
52.....	אימות דגם הביטוי של ה-mRNA-seq באמצעות Q-RT-PCR	4.2.2.3
57.....	השוואה בין דגם הביטוי לפי mRNA-seq ו-DGE	4.2.3
<b>59.....</b>	<b>5 דיון</b>	
59.....	דגם המתילציה	5.1
59.....	שוונות בדגם המתילציה של מקטעי DNA שבודדו באמצעות MS-AFLP	5.1.1
60.....	דגם ביטוי המתילציה של הגן DEF	5.1.2
62.....	אפיון דגם ביטוי באמצעות SOLEXA/ILLUMINA	5.2
63.....	השוואת האנליזות השונות ששימשו לאפיון דגם הביטוי	5.3
63.....	השוואה בין תוצאות ה-DGE לתוצאות ה-Q-RT-PCT	5.3.1
64.....	השוואה בין תוצאות mRNA-seq לתוצאות ה-Q-RT-PCT	5.3.2
65.....	השוואה בין תוצאות ה-DGE לתוצאות ה-mRNA-seq	5.3.3
65.....	גנים בהם נמצאו הבדלים ברמת הביטוי בין פרטים נורמאליים לחריגים	5.3.4
69.....	סיכום וכיוונים עתידיים	5.4
70.....	אימות תוצאות וגישות לחקר המנגנון האחראי ליצירת השונות הסומקלונלית	5.4.1
71.....	חשיבות הקלאית	5.5
<b>72.....</b>	<b>6 רשימת ספרות</b>	
<b>i.....</b>	<b>Summary</b>	

## רשימת טבלאות

- טבלה 1 - מקטעים הנבדלים בין עצים נורמאליים וחריגים באנליזה של MS-AFLP (גורביץ, 2004)..... 14
- טבלה 2 - עצים ששימשו לאופיין דגם המתילציה בשיטת ה-Bisulfite sequencing..... 17
- טבלה 3 - תחלים ששימשו לבידוד הגן DEFICIENS (DEF)..... 18
- טבלה 4 - תחלים ששימשו ל-Bisulfite sequencing..... 20
- טבלה 5 - עצי 'מג'הול' ששימשו לאנליזות הריצוף בנפח גבוה בשיטת Solexa..... 22
- טבלה 6 - תגים שהוגברו ב-5' RACE..... 29
- טבלה 7 - תחלים ששימשו ל-Q-RT-PCR לבחינת תגים שאותרו ב-DGE..... 30
- טבלה 8 - תחלים ששימשו ל-Q-RT-PCR לבחינת תגים שאותרו ב-mRNA-seq..... 31
- טבלה 9 - דגמי מתילציה של ארבעה מקטעי ה-DNA שאופינו באנליזת MS-AFLP..... 36
- טבלה 10 - השוואת רצף ה-cDNA של הגן DEF בין התמר לאורגניזמים אחרים..... 43
- טבלה 11 - אפיון המתילציה בגן DEF..... 44
- טבלה 12 - תגים השונים בדגם הביטוי בין WT ל-DW..... 46
- טבלה 13 - תוצאות BLAST כללי של הרצפים שהוארכו ב-5' RACE..... 47
- טבלה 14 - תוצאות BLAST לגנום התמר של הרצפים שבודדו ב-5' RACE..... 48
- טבלה 15 - רמות ביטוי של גנים על פי DGE ועל פי Q-RT-PCR בעץ נורמאלי וננס..... 50
- טבלה 16 - מספר הרצפים שהתקבלו באנליזת mRNA-seq..... 51
- טבלה 17 - רצפים שנבדלו בדגם הביטוי בין WT ל-DW..... 52
- טבלה 18 - רמות ביטוי של גנים על פי mRNA-seq ו-Q-RT-PCR בתערובות של עצים נורמאליים וננסים..... 53
- טבלה 19 - ביטוי של ארבעה גנים בעצי WT ו-DW בודדים..... 56
- טבלה 20 - ערכי הקורלציה בין שלושת האנליזות בארבעת הגנים הנבחרים..... 56
- טבלה 21 - השוואה בין האנליזות DGE, mRNA-seq ו-Q-RT-PCR..... 57
- טבלה 22 - ערכי הקורלציה בין שלושת האנליזות: DGE, mRNA-seq ו-Q-RT-PCR..... 58

## רשימת איורים

- 24..... RNA - 1 איור מעץ תמר נורמאלי וננס
- 27..... Solexa/Illumina - 2 איור טכנולוגיית הריצוף של
- 34..... FWA בארבידופסיס - 3 איור אתרי מתילציה שמורים בגן
- 37..... Bis5 במקטע - 4 איור דגם המתילציה במקטע
- 38..... Bis6 במקטע - 5 איור דגם המתילציה במקטע
- 41..... DEFICIENS - 6 איור מבנה הגן
- 42..... DEF באורגניזמים שונים - 7 איור רצף אקסון מספר 1 מהגן
- 44..... Bisulfite sequencing - 8 איור אזורי מתילציה בגן DEF ותחלים שתוכננו לריאקציית
- 50..... Sol 8 - 9 איור כיול ואנליזה של Q-RT-PCT לגן
- 55..... mRNA-seq וה-Q-RT-PCT - 10 איור קורלציה בין תוצאות דגם הביטוי מאנליזות ה-

## תקציר

התמר הינו ענף מטעים בעל חשיבות כלכלית רבה באזורים המדבריים של המזרח התיכון וצפון אפריקה. ריבוי מסורתי של התמר נעשה באמצעות חוטרים, שמספרם בכל עץ מוגבל. הצורך בהרחבת מטעי התמרים בעולם הביא לפיתוח טכניקות לריבוי התמר בתרבית רקמה. ריבוי בתרביות רקמה מאפשר קבלת כמות גדולה של שתילים זהים בתכונותיהם לצמח האם, בעלי אחידות ועמידות למחלות נפוצות ולמזיקים. מירב העצים שרובו בתרביות רקמה הינם נורמאלים ואחידים, ואולם נוצרים גם פרטים בעלי פנוטיפים חריגים. עצים חריגים אלה גורמים נזק כלכלי קשה למגדלי התמר. תופעה זאת, המכונה שונות סומקלונלית (somaclonal variation), ידועה בצמחים רבים שמרובים בתרבית רקמה ומהווה בעיה חמורה במספר גידולים חקלאיים.

העבודה הנוכחית מתמקדת באפיון מולקולארי של שני פנוטיפים חריגים נפוצים בעצי תמר: פנוטיפ ה'ננסות' בעצים מזן 'מג'הול' ופנוטיפ 'חנטה לקויה' ו'ריבוי שחלות' בתמרים מזן 'ברהי'. התמקדנו באפיון דגמי המתילציה והביטוי בעצי תמר חריגים שמקורם בתרבית רקמה ובעצים נורמאליים. דגם המתילציה נבחן באמצעות שיטת ה-Bisulfite sequencing ודגם הביטוי נבחן באמצעות ריצוף בהספק גבוה בטכנולוגיית Solexa/Illumina ואומת בעזרת Q-RT-PCR.

דגם המתילציה נבחן במספר אתרים בגנום התמר: בארבעה מקטעי DNA גנומי, שבעבודה קודמת, באמצעות אנליזת MS-AFLP, הראו שונות באתר ה-*MspI* ובגן *DEFICIENS* (DEF) שבעבודה קודמת נמצא להיות בעל דפוס ביטוי שונה בעצי 'ברהי' נורמאלים בהשוואה לעצים בעלי 'חנטה לקויה' ו'ריבוי שחלות'. בעבודה זאת נמצא דגם מתילציה שונה בין עלים של עצים נורמאלים לבעלי מופע ננסי מזן 'מג'הול' בשני מקטעי DNA. לעומת זאת, דגם המתילציה בשני מקטעי DNA אחרים היה דומה בעצים נורמאלים ובעצים בעלי פנוטיפ 'חנטה לקויה' ו'ריבוי שחלות' בתמרים מזן 'ברהי'. רמת המתילציה שנמצאה במקטע של הגן DEF הייתה נמוכה מאוד ולא נמצאו הבדלים מאפיינים בדגם המתילציה בין עלים ותפרחות של עצי תמר נורמאלים לבעלי פנוטיפ של 'ריבוי שחלות'.

דגם הביטוי אופיין ברקמת "לב" של תמרים נורמאלים ותמרים בעלי מופע ננסי מהזן 'מג'הול'. האפיון נעשה בשתי אנליזות ריצוף בהספק גבוה של Solexa/Illumina: Digital Gene Expression (DGE) ו-mRNA-seq. אנליזת ה-DGE נעשתה על עץ נורמאלי אחד ועץ ננסי אחד. בשיטה זאת נוצרים תגים קצרים (21 או 27 בסיסים) שהרצף שלהם מתחיל באתרי קישור לאנזימי רסטרקציה מסוימים. באנליזות המבוססות על חיתוך בחמישה אנזימי רסטרקציה שונים: *HpaII*, *HpyCH4IV*, *HinPII*, *NlaIII* ו-*TaqI*, התקבלו חמישה דגמי ביטוי של עצים נורמאלים וחריגים. לאחר סינון לקבלת תגים שדגם הביטוי שלהם שונה בפקטור של לפחות פי שלושה בין עצים נורמאלים לחריגים, התקבלו 155, 36, 48 ו-172 תגים בעלי דפוס ביטוי שונה עבור האנזימים השונים בהתאמה. מאה תגים מתוכם נבחרו להארכה באמצעות RACE 5'. 24 מהרצפים שהוארכו ב-RACE 5' נבחנו באמצעות Q-RT-PCR. במרבית הרצפים, לא נמצאה התאמה בין רמת הביטוי שהתקבלה באנליזת ה-DGE לזאת שהתקבלה באנליזת ה-Q-RT-PCR. בה התקבל ברוב המקרים יחס ביטוי של קרוב ל-1.

באנליזת ה-mRNA-seq, ה-RNA עבר פרגמנטציה למקטעים קטנים ורוצפו 33 בסיסים משני צידיו של כל מקטע (paired-end). הרצפים עברו assembly על בסיס טיוטת גנום התמר. אנליזת ה-mRNA-seq נעשתה על תערובות של שני עצים נורמאלים ו-15 עצים ננסיים. תערובת העצים הנורמאלים הכילה בסה"כ 13,647,010



רצפים ותערובת העצים הננסים הכילה 14,443,018 רצפים שהכילו 450,351,330 ו-476,619,594 בסיסים בהתאמה. הרצפים מופו בתוכנת SOAP ל-17,032 גנים צפויים בטיטוט גנום התמר. לאחר סינון לקבלת רצפים הנבדלים בדגם הביטוי בפקטור של לפחות פי שלושה בין עצים נורמאליים לחריגים, נשארו כ-2,000 גנים בעלי דפוס ביטוי שונה. נבחרו 23 גנים לבחינה באמצעות Q-RT-PCR. 21 גנים מתוכם הראו הבדלים ברמת הביטוי בין עצים נורמאליים לחריגים בדומה לתוצאות אנליזת ה-mRNA-seq והקורלציה בין שתי האנליזות הייתה מובהקת (מקדם קורלציה 0.992). בנוסף, בארבעה גנים אנליזת Q-RT-PCR בוצעה על שישה פרטים בודדים של עצים חריגים ושני פרטים של עצים נורמאליים. גם בגנים אלה הייתה קורלציה גבוהה בין תוצאות ה-mRNA-seq לתוצאות ה-Q-RT-PCR (מקדם קורלציה 0.989).

בין גנים שאותרו באנליזת mRNA-seq נמצאו מספר גנים הנבדלים ברמת הביטוי שלהם בין עצים נורמאליים לננסיים הקשורים לצימוח, יצירה ובקרה של פעילות הורמונאלית, גנים המעורבים בהתארכות תאים בצמחים וגנים הקשורים למתילציה. מבין הגנים שאומתו ב-Q-RT-PCR נמצאו כאלה שיכולים להיות קשורים לפנוטיפ הננסות או להתארכות תאים וביניהם גנים במסלול הביוסניטזה של ההורמון ג'יברלין, לסינתזה ולחישת של ברסינוסטרואידים, גנים המבוקרים על ידי אוקסין, הגן לאקספנסין ופקטור השעתוק WRKY. מוטנטים בחלק מגנים אלה מביאים לננסות במיני צמחים שונים.

בעבודה הנוכחית זוהו מספר מקטעי DNA גנומי וגנים רבים בעלי הבדלים ברמת הביטוי ובדגם המתילציה בין פרטים נורמאליים לחריגים הנוצרים במהלך ריבוי בתרביות רקמה. המשך בחינת הגנים תוכל לתרום להבנת המנגנון המבקר יצירת עצים חריגים בתמר, ויתכן שתאפשר שיפור של פרוטוקול תרבית הרקמה ופיתוח כלים מולקולאריים לזיהוי מוקדם של השתילים החריגים כדי למנוע את יצירתם והפצתם במטעים. יחד עם זאת, יש להדגיש את הקושי באפיון המנגנון המולקולארי של תופעת השונות הנוצרת בתרבית, וזאת כתוצאה מהמגבלה של השיטות המולקולאריות הזמינות, והצורך באימות של הממצאים השונים. יתכן שהמשך המחקר יהיה רלוונטי גם לזיהוי טיפוסים חריגים במינים אחרים המרובים בתרבית רקמה ולהבנה טובה יותר של תופעת השונות הסומלוגנית בכלל.

לסיכום, הממצאים העיקריים בעבודה זאת הם:

1. נמצאו הבדלים בדגם המתילציה בין עצים נורמאליים לננסים בשני מקטעי DNA.
2. באנליזה גנומית של הטרנסקריפטום נמצאו כ-2,000 גנים בעלי הבדלים בדגם הביטוי בין פרטים נורמאליים לננסים. ב-21 גנים מתוכם, ההבדל ברמת הביטוי אומת באמצעות Q-RT-PCR.

# 1. רקע מדעי

## 1.1. עץ התמר והדרכים לריבוי

עץ התמר (*Phoenix dactylifera* L.) הינו צמח חד-פסיגי, דו-ביתי השייך למשפחת הדקליים (*Arecaceae*). גנום התמר הינו דיפלואידי ומכיל 36 כרומוזומים (Barrow, 1998). לתמר גזע יחיד שיכול להגיע לגובה של למעלה מ-25 מטר. גזע התמר תומך בנוף ומהווה מערכת הובלה ראשית של הצמח. קודקוד הצמיחה האפיקלי המצוי בראש הגזע הינו פקע טרמינלי יחיד הנקרא "לב התמר" והוא האחראי על יצירת כל העלים והתפרחות החדשות של העץ (ברנשטיין, 2004; Zaid, 1999; Chao and Krueger, 2007). באופן מסורתי ריבוי תמרים נעשה ווגטיבית מחוטרים. החוטר (offshoot) הוא ענף צדדי המתפתח מחיקי העלים התחתונים של עצי תמר צעירים. מספר החוטרים שונה בהתאם לזן (יחידים עד עשרות), (ברנשטיין, 2004; סטולר, 1977; Zaid, 1999). בצורת ריבוי זאת, השתילים הנוצרים דומים בתכונותיהם להוריהם (true to type). אולם, לריבוי מחוטרים חסרונות רבים וביניהם המספר הקטן של חוטרים המיוצרים ע"י העץ רק בשנות חייו המוקדמות, ניתוק החוטר מהעץ מותיר פצע בעץ, החוטרים עלולים לשאת גורמי מחלות ומזיקים, שיעור הקליטה בנטיעת חוטרים הינו פעמים רבות נמוך, הבדלים בגודל החוטרים ובמיקומם על העץ יביאו להבדלים בקצב ההתפתחות ולשונות במטע והטיפול והשינוע של החוטרים יקר. מגבלות אלה והצורך בהרחבת מטעי התמרים בעולם, הביאו לפיתוח של שיטות ריבוי תמרים בתרבות רקמה (ישראלי, 2004; Reuveni et al., 1972; Tisserat, 1979; Zaid, 1999; Chao and Krueger, 2007). השיטות מבוססות על בידוד רקמת תאים מפקע אמירי של עץ האם או מפקעים אמיריים של החוטר. פרוטוקולים לריבוי מתרבות רקמה מאיברים אחרים של הצמח, כמו למשל מתפרחות עובריות פותחו גם כן, אך אינם משמשים בדרך כלל לריבוי מסחרי (ישראלי, 2004). קיימות שתי שיטות עיקריות לריבוי תמרים בתרבות רקמה: אמבריוגנזה ואורגנוגנזה. בתמרים שיטות אלו פותחו בין השנים 1970-1980 (ישראלי, 2004; Chao and Krueger, 2007).

**א. אמבריוגנזה:** המקור ל-explant (הרקמה שנלקחת מהצמח) הראשוני בתמר הוא הפקע האמירי או פקעים חיקיים של חוטר. שימוש בפקע אמירי מחייב את הסרת כותרת העץ ומביא למות העץ. לכן השימוש השגרתי הוא בפקעים אמיריים מחוטרים (ישראלי, 2004; Tisserat, 1979). בתנאים סטריליים מצמצמים את גודל האקספלנט עד לקבלת המקטע המרכזי המכיל את קודקוד הצמיחה, תחילת של עלים ובסיסי עלים צעירים. אקספלנט זה נשתל על מצע מתאים בחושך עד להתפתחות קאלוס אמבריוגני ואז מועבר למצע ללא חומרי צמיחה ובאור. בתנאים אלה העוברים מתפתחים לצמחונים ומועברים להשרשה. התהליך כולו עד לקבלת שתיל ראוי לנטיעה נמשך בין 4-5 שנים (ישראלי, 2004; Bhaskaran and Smith, 1992; Reuveni et al., 1972).

**ב. אורגנוגנזה:** המקור לאקספלנט הם פקעים וגטיביים מחוטרים הכוללים מריסטמה קודקודית, פרימורדיות עלים ובסיסי עלים צעירים. בשיטה זאת אין מעבר דרך שלב הקאלוס אלא חלה התפתחות ישירה של רקמות קיימות. בתנאים מתאימים הרקמות מתפתחות ונוצרות עליהן בליטות בצבע ירקרק

דמויות עלים זעירים. נוצרים צמחונים אותם מעבירים לכלי עם מצע מעודד השתרשות. כאשר התהליך מתקדם באופן תקין אפשר להפיק מחוטר אחד עד 3000 צמחים (ישראלי, 2004; Zaid, 1999).  
לריבוי תמרים בתרביות רקמה יש יתרונות כדלקמן: ניתן לקבל כמות גדולה של שתילים הזהים בתכונותיהם לצמח האם, השתילים מצטיינים באחידות רבה בגודל ובקצב הצימוח ונקיים מגורמי מחלות ומזיקים (ישראלי, 2004).

## 1.2. שונות סומקלונלית

ריבוי בתרביות רקמה מביא בדרך כלל ליצירת שתילים זהים בתכונותיהם לצמח האם (true-to-type), אולם לעיתים נוצרים בתהליך זה גם טיפוסים חריגים (off-types) השונים פנוטיפית מצמח האם. תופעה זאת של יצירת טיפוסים חריגים בתרבית רקמה נקראת שונות סומקלונלית - Somaclonal Variation (Kaepler et al., 1981; Larkin and Scowcroft, 2000; et al.). שונות סומקלונלית קיימת במינים שונים המרובים בתרביות רקמה כמו דקל השמן, בננה ותמר (Larkin and Scowcroft, 1981; Sahijram et al., 2003). יתר על כן, במינים שונים, הריבוי בתרבית מביא ליצירת פנוטיפים דומים.

## 1.3. המנגנון האחראי ליצירת שונות סומקלונלית

הסיבות ליצירת השונות הסומקלונלית אינן ידועות. מניחים שתנאי תרבית הרקמה הם הגורמים לשינויים. בין הגורמים המשפיעים על התופעה ראוי למנות את אופי האקספולנט, ההרכב והריכוז של ההורמונים במצעי הגידול ומספר ה"גילגולים" (ההעברות) שעובר החומר הצמחי בתרבית (Sahijram et al., 2003; Evans, 1984; et al.). נמצא שמרווחים גדולים בין ההעברות של החומר הצמחי ויחס גבוה בין ריכוז ההורמונים אוקסין וציטוקינין, מצמצמים את השכיחות ליצירת הפנוטיפים החריגים. לעומת זאת ריכוזים גבוהים של ההורמון ציטוקינין עלולים לעודד יצירת שונות סומקלונלית (Damasco et al., 1998; Kunert et al., 2002; Eeuwens et al., 2003).

שונות סומקלונלית יכולה לנבוע משינויים גנטיים ו/או אפיגנטיים. שינויים גנטיים כוללים שינוי במספר הכרומוזומים, אברציות כרומוזומליות (טרנסלוקציות, חסרים, אינברסיות ודופליקציות) וכן שינויים ברצף ה-DNA. שינויים אפיגנטיים עשויים להביא להשתקה או אקטיבציה של גנים על ידי שינויים בהטרוכרומטין, ומתלציה של ה-DNA (Kaepler et al., 2000). בנוסף, יש לציין שמתלציה עשויה לגרום למוטציות על ידי הפעלה של טרנספוזונים אשר עשויים לגרום לשבירת כרומוזומים ושינויים ברצף ה-DNA (Phillips et al., 1994). יש הסוברים שכתוצאה מהגידול בתרבית יש שינויים ברמת המתלציה בכלל הגנום. מחקרים הראו שדגמי המתלציה עלולים להשתנות בעקבות הריבוי בתרביות רקמה (Kaepler et al., 2000).  
שינויים סומקלונליים מתרחשים הן באתרים אקראיים והן בתדירות רבה באזורים הנקראים 'hot spot' (Oh et al., 2007; Linacero et al., 2000; Phillips et al., 1994). במספר מחקרים ניסו למצוא הבדלי רצף באזורים אלה בין פרטים נורמאליים לחריגים באמצעות סמני DNA. מכיוון שקיים שימור רב ברצף ה-DNA של דקל השמן והתמר, אפיון 'hot spot' בדקל השמן יסיע לאיתורם בתמר (Johnson et al.,

Linacero וחוב' (2000) מצאו שצמחי שיפון שונים שרובו בתרבויות הכילו שיעור גבוה של מוטציות באזורים מסוימים (אזורים היפרוריאבילים), דבר המראה שאירועי מוטציות התרחשו באותו הרצף בצורה בלתי תלויה בצמחים שונים (Linacero et al., 2000).

#### 1.4. שונות סומקלוגלית וטיפוסים חריגים בתמר

רוב הטיפוסים החריגים שנוצרים בתמרים נחותים בתכונותיהם לעומת הוריהם. הפנוטיפים החריגים הנפוצים בתמר כוללים צמחים בעלי מורפולוגיה ומבנה א-נורמאליים כמו תפרחות מעוותות, עלים רחבים, גדילה ווגטטיבית מוגזמת, עלים מלבינים, ננסות, חוטרים מעוותים, עיכוב בזמן פריחה, כשל בחנטה ופירות מעוותים (Al Kaabi et al., 2007a; McCubbin et al., 2000; Chao and Krueger, 2007; Cohen et al., 2007). כשליש מעצי התמר שרובו בתרבויות רקמה ונשתלו בישראל בשנים 1993-2001 היו בעלי מופע חריג (Gurevich et al., 2005). בעבודה זאת התמקדנו בבחינה של שני פנוטיפים חריגים הנפוצים מאוד בין עצי התמר שמרובים בתרבות רקמה בארץ ובעולם:

**א. 'עיכוב צימוח וננסות':** במטעי 'מג'הול' (וזנים אחרים) שנוצרו מריבוי בתרבויות רקמה נמצאו עצים בעלי מופע ננסי אופייני (גובה הגזע שלהם הגיע לפחות ממטר אחד בגיל ארבע עד חמש שנים במטע לעומת גובה ממוצע של שני מטר בעצים מקבילים), בעלי עלים קצרים והוצים ארוכים ורחבים. בנוסף, נוף הצמח נראה מעוגל ומכונס והפירות מעוותים. הננסות מחלישה את העצים ומפחיתה את ייצור החוטרים. בסקר שנערך בדרום אפריקה שיעורם הגיע עד 4% (McCubbin et al., 2000; Al Kaabi et al., 2007a). בארץ נמצאו כ-10,000 עצים עם מופע ננסי, שמקורם ממעבדת ריבוי אחת ועצים ממקורות אחרים המופיעים בתדירות משתנה. (ברנשטיין, 2004; Gurevich et al., 2005).

**ב. עצים בעלי 'חנטה לקויה':** עצים בעלי 'חנטה לקויה' מהזנים 'ברהי' ו'חאלס' דווחו לראשונה בשנת 2000 על ידי ג'רבי (Djerbi, 2000; Al-Wasel, 2001). מדובר בכ-100,000 עצים שהובאו לערב הסעודית בשנים 1994-1995 ו-1999. בהגיעם לגיל 4-5 שנים נרשמה בהם חנטה של 0-20% בלבד. בפרחים שלא חנטו התגלתה גם תופעה של 'ריבוי שחלות'. 'חנטה לקויה' (עד העדר חנטה) של הזן 'ברהי' מריבוי באמבריוגנזה דווחה גם בדרום אפריקה (McCubbin et al., 2000) ובמקומות נוספים (Al Kaabi et al., 2007a). מתוך כ-3,500 עצי 'ברהי' ו'חאלס' מתרבות רקמה שניטעו בישראל בין השנים 1993-2000, כ-1,500-2,000 סבלו מבעיות חנטה. חוסר החנטה מלווה בתופעות של 'ריבוי שחלות' ובעיוותים בפרחים. נמצא שחומרת הפנוטיפ פוחתת ושיעור החנטה משתפר עם התבגרותו של העץ (כהן וחוב', 2003; Cohen et al., 2004).

נמצא שבזני התמר 'סוקרי', 'ברהי', 'סולטנא' (Sultana), 'חאלס', 'אומ-דן' (Oum Dahn) ו'מג'הול' קיים פנוטיפ ננסות (Al Kaabi et al., 2007a). פנוטיפ זה נפוץ מאוד גם בבנות, בהם עד 80% מהפרטים החריגים שמתקבלים בריבוי בתרבויות רקמה הם בעלי פנוטיפ ננסי (Bairu et al., 2006; Smith and Drew, 1990; Reuveni and Israeli, 1990). פנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות' בתמרים זוהה בארצות רבות בעולם. בתמרים מהזן 'ברהי' נמצא אחוז גבוה של עצים בעלי פנוטיפ זה. בזנים אחרים כמו

'הלאס', 'סוקאלי', 'Ajoua' ו'דקל נור' פנוטיפ זה מצוי בשכיחות נמוכה יותר (Djerbi, 2000; Al-Wasel, 2001). בעצי דקל השמן בולט פנוטיפ הדומה מאוד לזה של 'חוסר החנטה' ו'ריבוי השחלות' המכונה 'Mantled' (Corley et al., 1986). כ-5% מעצי דקל השמן שרובו בשיטת האמבריוגנזה היו בעלי פנוטיפ 'Mantled'. פנוטיפ זה מתאפיין בשיעורי חנטה נמוכים מאוד וב'ריבוי שחלות'. ברוב העצים התופעה מתמתנת במהלך ההתבגרות, ולאחר מספר שנים במטע העצים הופכים לנורמאליים (Rival et al., 1998b). עבודה זאת מתמקדת באפיון מולקולארי של פנוטיפ 'הננסות' וה'חנטה הלקויה' בתמרים תוך התמקדות באפיון מולקולארי של תופעת הננסות בזן 'מג'הול' ובבחינת הבדלים בדגם המתילציה בתמרים משני הפנוטיפים החריגים 'חנטה לקויה' ו'ריבוי שחלות' בזן 'ברה' וננסות בזן 'מג'הול'.

## 1.5. שיטות מולקולאריות לבחינת שונות סומקלונלית

שיטות מולקולאריות שונות פותחו לזיהוי הבדלים בין מינים, זנים ופרטים שונים. שיטות אלה שימשו גם לבחינת השונות הסומקלונלית בין צמחים נורמאליים וחריגים שרובו בתרביות רקמה. בעבר היה שימוש רב באיזואנזימים אולם היום רוב השיטות מתמקדות בזיהוי שינויים ברצף ה-DNA. שימוש בסמנים בדקל התמר התמקד באפיון השונות הגנטית בין זנים שונים. רק מחקרים מעטים עסקו באפיון השונות הסומקלונלית של עצים שרובו בתרביות רקמה. נמצאו הבדלים ברצף ה-DNA ואולם לא נמצאו סמנים ייחודיים המאפשרים זיהוי של המופעים החריגים. להלן מפורטות השיטות המחקריות העיקריות לבחינת שונות גנטית והשימוש שנעשה בהם לאפיון השונות הסומקלונלית בתמר:

**איזואנזימים:** חלבונים בעלי תפקיד אנזימטי דומה אבל בעלי מטען שונה. ניתן להפריד בין איזואנזימים שונים על גבי ג'ל עמילן בהרצה בשדה חשמלי ולזהות אותם באמצעות פעילותם האנזימטית. האיזואנזימים אשר שימשו למחקר בתמרים הם: Esterase (EST), Glutamate oxaloacetate transaminase (GOT), Peroxidase (PER), Alcolhol dehydrogenase (ADH), Endopeptidase (EPN) ו-Polyphenol oxidase (POD) (Kunert et al., 2003). בעבודה אחת לא נמצאו הבדלים בדגם הפסים של האיזואנזימים בין עצים נורמאליים לעצי off-type שנוצרו בתרבית רקמה (Salman et al., 1988). בעבודה אחרת נמצאה שונות בעוצמת הפעילות של האיזואנזימים בין עצי תמר נורמאליים לעצי off-type שמקורם מתרבית רקמה. נמצאה פעילות חזקה של האיזואנזימים GOT ו-PER בפרטים נורמאליים לעומת זאת בפרטי off-type נמצאה פעילות חזקה של האיזואנזים POD (Saker et al., 2000).

**אנליזת RFLP (Restriction Fragment Length Polymorphism):** השיטה מאפשרת לזהות נוכחות או העדר אתר רסטריקציה במקומות שונים לאורך הגנום. בשיטה זאת נמצא פולימורפיזם בין חמישה זני תמרים: 'ברה', 'דקל נור', 'מג'הול', 'הלאס' ו'חדראווי'. לא נמצאה שונות בין פרטים נורמאליים לחריגים, אבל נמצאה שונות בין חוטרים לקאלוס שנובעת כנראה משונות סומקלונלית (Corniquel and Mercier, 1994).

**אנליזת RAPD (Random amplified polymorphic DNA):** אנליזה המבוססת על שינויי רצף המשפיעים על קשירת תחל קצר לתבנית ה-DNA (Williams et al., 1990; Welsh and McClelland, 1990).

1990). אנליזה זאת בתמרים הראתה שונות סומקלונלית בכ-4% מהפרטים שנבדקו (70 פרטים) מזן 'זגלול' שרובו בתרביות רקמה (Saker *et al.*, 2000). כמו כן נמצאה גם שונות סומקלונלית בזן 'ברהי' בין תמרים שרובו בתרביות רקמה לאלו שרובו מחוטרים (Ali *et al.*, 2007). במחקר אחר לא נמצאו עדויות לקיום של שונות סומקלונלית (Saker *et al.*, 2006). בעבודה נוספת שבחנה יציבות גנטית בתמרים נמצאה שונות בתוך צמחים שרובו בתרביות רקמה ובינם לבין צמח האם שלהם ב-25 פסים מתוך 57 פסים (Eshraghi *et al.*, 2005).

אנליזה RAPD של הפנוטיפ 'Mantled' בדקל השמן לא הראתה הבדלים ברצף ה-DNA בין פרטים נורמאליים לבעלי פנוטיפ 'Mantled' (Rival *et al.*, 1997, 1998a, b). Damasco וחוב' (1996) זיהו בבנות סמני RAPD ייחודיים של פנוטיפ ננסות. זוהה סמן המופיע באופן קבוע בכל הפרטים הנורמאליים וחסר בכל הפרטים החריגים.

**SSR (Simple Sequence Repeat)**: השיטה מזהה הבדלים בין פרטים במספר החזרות של רצפים קצרים של מיקרוסטטיליטים המוגברים ב-PCR באמצעות תחלים ייחודיים התוחמים את הרצף החוזר (Gupta *et al.*, 1994). השיטה זאת נמצאה שונות גנטית גבוהה בין פרטים של דקל התמר בתוניס (Zehdi *et al.*, 2005). במחקר אחר נמצא פולימורפיזם בין 14 זנים שונים של תמרים (Hamama *et al.*, 2003), אולם לא מוכרות לנו עבודות המשוות בשיטה זו טיפוסים חריגים עם נורמאליים בתמרים.

**AFLP (Amplified fragment length polymorphism)**: השיטה מבוססת על זיהוי אתרי רסטרקציה. הזיהוי נעשה לאחר הגברת המקטעים החתוכים ב-PCR באמצעות שימוש באדפטורים ותחלים מתאימים. בשיטה זאת נוצר דגם פסים ייחודי לכל פרט (Vos *et al.*, 1995). בתמרים השתמשו בשיטה זאת לצורך איתור שונות סומקלונלית שנוצרת במהלך הריבוי בתרביות רקמה (אמבריוגנזה ואורגנוגנזה). בכל עשרת הצמחים שנבדקו נמצאה שונות סומקלונלית. שכיחות השונות הסומקלונלית בצמחים שרובו בשיטת האורגנוגנזה הייתה נמוכה יותר בהשוואה לאלה שרובו בשיטת האמבריוגנזה. יתר על כן, בשיטת האורגנוגנזה התקבלו צמחים שקרובים יותר גנטית לצמח האם (Al Kaabi *et al.*, 2007b). בעבודה אחרת נמצאה שונות סומקלונלית נמוכה בזנים 'סקוטי' (Sakkoty), 'גנדילה' (Gandila) ו'ברטמודה' (Bertamoda) ברמה של 2.6%, 0.79% ו-1% בהתאמה (Saker *et al.*, 2006). בבחינה של שני הזנים העיקריים בארץ 'מג'הול' ו'ברהי' נמצא דגם ייחודי לכל זן. בנוסף נמצאה שונות גנטית גבוהה בתוך פרטים מהזן 'מג'הול' שרובו בתרבית רקמה ושונות נמוכה בתוך עצי 'ברהי' חריגים מתרבית (Gurevich *et al.*, 2005).

**RDA (Representational Difference Analysis)**: השיטה מזהה הבדלים גנטיים על סמך התאמה של מקטעי DNA גנומי או cDNA שמקורם בפרטים בעלי הומולוגיה גבוהה. בשיטה זאת ה-DNA שמופק מהפרטים נחתך על ידי אנזימי רסטרקציה, נעשה חיבור של מתאמים לקצוות והגברה של המקטעים ב-PCR. המקטעים שמקורם מפרט אחד משמשים כ-Tester, ואילו המקטעים שמקורם מפרט אחר משמשים כ-Driver. נעשית היברידיזציה בין Tester ל-Driver בעודף גדול של Driver. השיטה מאפשרת זיהוי מקטעים ייחודיים של ה-Tester (Lisitsyn *et al.*, 1993). שיטה זאת שימשה לבחינת הפולימורפיזם בין בנות נורמאליות לחריגות שמקורם בתרבית רקמה. נמצא פולימורפיזם בשיעור גבוה בין פרטים נורמאליים לחריגים, ושונות של כ-10% ברצף הנוקליאוטידים במקטע מסוים. בהשוואה בין פרטים נורמאליים לבעלי

פנוטיפ 'Mantled' בדקל השמן, נמצאה שונות גבוהה בין שני הפנוטיפים באזורים שונים בגנום. הבדיקה התבססה על ארבעה זוגות תחלים באזורים שונים בגנום הבננה בהם נמצאה שונות סומקלונלית. בארבעה תחלים אלה נמצאה גם שונות סומקלונלית בדקל השמן בהשוואת פרטים נורמאליים לבעלי פנוטיפ 'Mantled' (Cullis *et al.*, 2007). תוצאות אלה תומכות בהיפותזה שקיימים אזורי 'hot spot' שרגישים במיוחד ל"עקה" במהלך הריבוי בתרביות רקמה (Oh *et al.*, 2007). כנראה שאזורים אלה מהווים hot spots במגוון מינים. באזורים אלה בגנום התמרים נמצאה שונות גנטית רבה בין זנים שונים (Johnson *et al.*, 2009), אולם עדיין לא נמצא הקשר בין פנוטיפים חריגים ואתרים אלה בגנום התמר.

## 1.6. אפיון מתילציה

מתילציה של DNA מתרחשת בחיידקים, צמחים ובעלי חיים. המתילציה נעשית על פחמן מספר 5 בציטוזין והופכת אותו ל-5-methylcytosine. מודיפיקציה זאת לא משנה את רצף ה-DNA באורגניזם. המתילציה של ציטוזין מתרחשת לאחר שכפול ה-DNA באמצעות האנזים cytosine methyltransferase ואנזימים נוספים (Adams, 1995; Richards, 1997; Finnegan *et al.*, 1998). בגנום הצמחי עד כ-30% מהציטוזינים יכולים להיות ממותלים. רצפי המטרה העיקריים למתילציה בצמחים הם CG ו-CNG. (Gruenbaum *et al.*, 1981), אבל מתילציה יכולה להתרחש גם ברצף DNA לא סימטרי CNN (Pelissier *et al.*, 1999). בצמחים נמצאה כמות גדולה של רצפי CG ממותלים במעלה (upstream) קודון תחילת התרגום ויחסית מעט אתרי CG ממותלים באזור המקודד של הגן (Shimizu *et al.*, 1997). הרמה הנמוכה ביותר של רצפי CG ממותלים נמצאה באזורים לא מקודדים (Messegueur *et al.*, 1991).

קיימות עדויות רבות לכך שמתילציה מעורבת בבקרה של ביטוי גנים. המתילציה יכולה להשפיע על האינטראקציה בין DNA לחלבון. באאוקריוטים, מתילציה של ציטוזין יכולה לעכב קישור של חלבונים רגולטורים (Ehrlich and Ehrlich, 1993) ומתילציה באזור הפרומוטור או ברצף המקודד יכולה להשפיע על ביטוי (Jost and Saluz, 1993; Lewis and Bird, 1991). לא כל הרצפים הממותלים מושקפים, וחלקם בא לידי ביטוי (Hafiz *et al.*, 2001). למשל, רמת הביטוי של הגן  $\beta$ -glucuronidase באורז לא תלויה בשיעור המתילציה בגן זה (Wang and Waterhouse, 2000).

לרמת המתילציה יש השפעה על התפתחות הפרה ועל פוריותו. עיכוב פעילות המתילאז בצמח המודל ארבידופסיס השפיעה על מבנה הצמח והפרה (Finnegan *et al.*, 1996). שינויים ברמת המתילציה נמצאו כגורמים לעיוותים בפרחים, לתוספת אברים בפרח (אבקנים או שחלות), ולעקרות (Richards, 1997; Kaeppler, 2000). שינויים בדגם המתילציה מהווים את אחד המנגנונים ליצירת שונות סומקלונלית אפיגנטית (Kaeppler *et al.*, 2000). נמצא שרמות מתילציה כלליות ומתילציה באתרים מסוימים נובעים משינויים שחלים במהלך הריבוי בתרביות רקמה (Rival *et al.*, 2009). בדקל השמן נמצאה ירידה במתילציה בשיעור של 0.5-2.5% בהשוואת פרטים בעלי פנוטיפ 'Mantled' לפרטים נורמאליים (Jaligot *et al.*, 2000). נמצאו הבדלים בדגמי המתילציה בין עצי 'Mantled' לעצים נורמאליים מאותו מקור גנטי (clone), אך אלה אינם מאפיינים את כל העצים החריגים מקלונים אחרים (Jaligot *et al.*, 2004). מכאן שפנוטיפ

ה-'Mantled' קשור לשונות אפיגנטית המתבטאת ברמה ובדפוס המתילציה של ה-DNA (ולא ברצף הבסיסים). בדקל השמן נמצאה עליה ברמת השעתוק של שנים מתוך שלושה גנים המקודדים למתילטרנספראזות (DNMT) בקלוס בעל נטייה ליצור פרטים עם פנוטיפ 'Mantled' לעומת קלוס היוצר בדרך כלל פרטים נורמאליים. תוצאות אלה נראות מנוגדות לתופעת ההיפומתילציה בפרטי ה-'Mantled' (Rival *et al.*, 2008; Rival *et al.*, 2009). בבחינת דגם המתילציה בבננה באמצעות האיזוסכיזומרים *HpaII*-ו-*MspI*, נמצא שקיים הבדל במתילציה במקטע מסוים. במקטע זה קיימת מתילציה בעצי בננה ננסים ולא קיימת בעצי בננה נורמאליים. ממצא זה הינו בנוסף לשינויים ברצף ה-DNA בין פרטים אלה (Oh *et al.*, 2007).

Methylation sensitive AFLP (MS-AFLP) הינו כלי מולקולארי המאפשר לאפיין את דגם המתילציה (Mingliang *et al.*, 2000). בעבודת הגמר של ס. גורביץ (גורביץ, 2004, עבודת גמר) נעשה שימוש בשיטה זאת לצורך בחינת דגם המתילציה בעצי תמר. האנליזה נעשתה על 15 עצי 'ברהי' ו-13 עצי 'מג'הול' שרובו מחוטרים ובתרביות רקמה, בעלי פנוטיפ נורמאלי וחריג. נמצאו תשעה פסים (טבלה 1) בעלי דגם מתילציה שונה בהשוואה בין העצים החריגים והנורמאליים. בזן 'ברהי' נמצאו שלושה פסים פולימורפיים, המבדילים בין עצים נורמאליים לעצים בעלי פנוטיפ 'חנטה לקויה' ו'ריבוי שחלות'. שניים מתוכם מראים דגם מתילציה שונה (פס אחד ייצג מתילציה והשני ייצג המי-מתילציה בעצים נורמאליים שאינה קיימת בעצים החריגים). הפס שלישי מציג כנראה הבדל ברצף הבסיסים בין עצים נורמאליים לחריגים. בזן 'מג'הול' נמצאו שישה פסים פולימורפיים, המבדילים בין עצים נורמאליים לעצים בעלי פנוטיפ ננסות. פס אחד מייצג מתילציה בעצים נורמאליים ושינוי ברצף הבסיסים בפנוטיפ החריג, פס נוסף מייצג אתר ממותל בעצים הנורמאליים ואתר לא ממותל בפנוטיפ החריג. ארבעה פסים נוספים הראו שונות ברצף בין שני הפנוטיפים. תוצאות אלה מציעות שהבדלים בדגם המתילציה בנוסף להבדלים ברצף הבסיסים, קשורים לתהליכי יצירת השונות הגנטית של הפנוטיפים החריגים הנפוצים בתמר.



טבלה 1 - מקטעים הנבדלים בין עצים נורמאליים וחרגיגים באנליזה של MS-AFLP (גורביץ, 2004)

אורך הפרגמנט (bp)	off-types		WT		זן	שם הפרגמנט
	חיתוך ב- <i>EcoRI</i> <i>HpaII</i> -ו	חיתוך ב- <i>EcoRI</i> <i>MspI</i> -ו	חיתוך ב- <i>EcoRI</i> <i>HpaII</i> -ו	חיתוך ב- <i>EcoRI</i> <i>MspI</i> -ו		
115	+	+	-	-	ברהי	1 (Bis 1)
178	+	+	+	-	ברהי	2 (Bis 2)
לא ידוע	+	+	-	+	ברהי	3
לא ידוע	-	-	+	+	מג'הול	4
145	-	-	+	+	מג'הול	5 (Bis 5)
523	+	+	-	-	מג'הול	6 (Bis 6)
260	-	-	-	+	מג'הול	7
139	+	+	-	+	מג'הול	8
לא ידוע	+	+	-	-	מג'הול	9

טבלה המציגה מקטעי DNA שנמצאו בעלי הבדלים ברמת המתילציה בין עצים נורמאליים לחרגיגים. *MspI* ו-*HpaII* הינם שני איזוסכיזומרים המזהים את אתר הרסטריקציה CCGG, אולם הם בעלי רגישות שונה למתילציה של האתר. *MspI* אינו רגיש למתילציה וחותרך DNA ממותל ואילו אנזים רסטריקציה *HpaII* הינו רגיש למתילציה ולכן איננו חותרך DNA ממותל. השם Bis ליד חלק מהמקטעים מסמל פרמנט DNA שנבחן בהמשך באמצעות Bisulfite sequencing.

בעבודה של מ. שוחט (שוחט, 2007, עבודת גמר) בודדו גנים מבקרי פריחה בתמר, ונבחן הקשר שלהם לפנוטיפ 'חנטה לקויה' ו'ריבוי שחלות' בפירות 'ברהי'. אחד הגנים הנבחנים היה DEFICIENS (DEF). הגן מכיל MADS box ממשפחת B. האורתולוג שלו באורז נקרא SUPERWOMAN1 (SPW1), ובארבידופסיס 3 APETALA (AP3). מוטציה בגן באורז גורמת להחלפה הומיאוטית של אבקנים וקשקשי תפיחה לשחלות ולמבנים דמויי מוץ עליון בהתאמה (Nagasawa et al., 2003; Bommert et al., 2005). בדקל השמן נמצאו הבדלים ברמת הביטוי בין פרטים נורמאליים לבעלי פנוטיפ 'Mantled' של מספר גנים מסוג MADS box (Adam et al., 2007). רמת הביטוי של הגן DEF בעצי תמר נורמאליים לעומת עצים בעלי פנוטיפ 'חמור' או "חלש" של 'ריבוי שחלות' ו'חנטה לקויה' נבדקה ב-Q-RT-PCR. הפנוטיפ "החמור" הוגדר כעצים בעלי 'חנטה לקויה' המתבטאת בכ-30-65% פירות פרתנוקרפיים בהם מתפתחות כל שלוש השחלות ומעל-15% פירות 'מרובי שחלות'. הפנוטיפ "החלש" הוגדר כעצים בעלי 'חנטה לקויה' המתבטאת בכ-30-65% פירות פרתנוקרפיים בהם מתפתחות כל שלוש השחלות וכ-0-5% חנטים 'מרובי שחלות'. נמצא שביטוי הגן בעצים החרגיגים בעלי פנוטיפ "חמור" היה נמוך פי שלושה לעומת העצים הנורמאליים. לעומת זאת בפרטים בעלי הפנוטיפ ה"חלש" לא נמצא הבדל משמעותי ברמת הביטוי בין עצים נורמאליים לחרגיגים (שוחט, 2007, עבודת גמר).

נעשתה בחינה ראשונית של דגם המתילציה של הגן DEF בעץ 'ברהי' נורמאלי ובעץ בעל פנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות'. הבחינה נעשה ע"י חיתוך DNA באנזימים *HpaII* ו-*MspI* המזהים את אותו אתר חיתוך אך בעלי רגישות שונה למתילציה. ב-DNA מהעץ החרגי שנחתך ב-*MspI* התקבל לעיתים תוצר חלש מאוד שיתכן ונבע ממתילציה של אתר ה-*MspI/HpaII* ברמה נמוכה (שוחט, 2007, עבודת גמר).

## 1.7. אפיון דגם ביטוי גנים

דגם ביטוי הגנים משתנה בין פרטים ובין רקמות ותאים שונים בתוך אותו פרט. פותחו טכנולוגיות שונות המאפשרות לחקור את דגם הביטוי. קיימות שיטות המתבססות על ביטוי גנים בודדים ביחס לגן מנרמל (reference) כמו Northern blot ו-Q-RT-PCR. כמו כן קיימות שיטות גנומיות המאפשרות לאפיין את דגם הביטוי של גנים רבים, כמו DNA microarray ו-High-throughput sequencing. השיטות השונות מאפשרות לבחון את דגם ביטוי גנים בין רקמות, פרטים, שלבים שונים בהתפתחות, רמות שונות של עקה ועוד.

בשיטת ה-Q-RT-PCR דגם הביטוי מתקבל בעזרת מולקולות פלואורסנטיות שונות כמו SYBR Green, או על ידי שימוש בגלאי פלואורסנטי בשיטת TaqMan המאפשרות כימות של תוצרי ההגברה (Higuchi *et al.*, 1993).

שיטת ה-DNA microarray מבוססת על שבב זכוכית או סיליקון, שאליו מחוברים גלאים בקשרים קוולנטיים. ה-cDNA עובר היברידיזציה לגלאים אלה. החיסרון המשמעותי בשיטה היא מוגבלותה לאורגניזמים בהם יש מידע על רצף ה-DNA, שכן ייצור שבב ה-DNA מחייב מידע על הגנום הנבדק. כמו כן מכיוון שאנליזה זאת מסתמכת על אנוטציות קיימות, נמנע זיהוי של יחידות תעתיק קטנות חדשות (Busch and Lohmann, 2007).

בשנים האחרונות מתרחשת מהפכה בשיטות הרצוף המתבטאת בריצוף DNA בהספק גבוה (High-throughput sequencing). שיטות אלו מבוססות על הגברה של DNA משבטים (clones) רבים ושונים במקביל. טכנולוגיות אלה מבטיחות ריצוף בתפוקה גבוהה והפחתה ניכרת בעלות הריצוף (Hall, 2007; Busch and Lohmann, 2007). שיטות אלה המאפשרות ריצוף של מאות אלפים עד מיליוני רצפים במקביל ומשמשות גם לבדיקת דגם הביטוי. החברות המובילות בטכנולוגיה זאת הן 454 של Roche, SOLID של Applied Biosystems ו-Solexa של Illumina ([www.illumina.com](http://www.illumina.com)), [www.454.com](http://www.454.com), [www3.appliedbiosystems.com](http://www3.appliedbiosystems.com)). בכל השיטות הללו נעשה ריצוף בהספק גבוה על משטח מוצק. היתרון בשיטות אלה הוא שאין צורך בידע מוקדם על גנום האורגניזם הנבדק (Collins *et al.*, 2008; Marguerat *et al.*, 2008).

ריצוף בשיטת 454 מבוסס על הגברה של מקטעי DNA על "חרוזים" (beads) בתוך טיפות תחליב. לאחר מכן ה-beads מוטענים לבארות על משטח פיברו-אופטי (fiberoptic). ההגברה נעשית במחזורים שבכל אחד משתתף בסיס אחד. אינקורפורציה של הבסיס למקטע המסונתז מלווה בפליטת אור המזוהה בריאקציה כימית המבוססת על לוציפראז (luciferase) (Margulies *et al.*, 2005). השיטה מאפשרת ריצוף של מאות בסיסים מכל מקטע (Margulies *et al.*, 2008; Bentley, 2006).

הטכנולוגיה של Solexa/Illumina מבוצעת בשני שלבים: בשלב הראשון הגברה של ה-cDNA ובשלב השני ריצוף. הריצוף מתבצע על גבי משטח מיוחד המאפשר גישה לאנזימים ויציבות גבוהה לתבנית. הריצוף מבוסס על כך שלכל אחד מהבסיסים יש צבע פלואורסנטי אחר. בשיטה זאת ניתן להגיע לצפיפות של עד 10 מיליון מולקולות זהות באשכול (cluster) לסמ"ק. בשיטה זאת ניתן לבצע ריצוף של מקטעי DNA באורך

של עד 75 בסיסים. לאחרונה פותחו פרוטוקולים המאפשרים להגיע לרצפים ארוכים של עד 120 בסיסים ויותר ([www.illumina.com](http://www.illumina.com)). מהרצפים הקצרים שמתקבלים ניתן להרכיב קונטיגים ארוכים יותר (Margulies *et al.*, 2005).

הרצפים הקצרים שמתקבלים מאנליזת Solexa יכולים לשמש להשוואה בין כמה גנומים בעלי הומולוגיה גבוהה לצורך זיהוי הבדלים ברצף כמו SNPs (Polymorphisms Single Nucleotide), ביטוי גני small RNAs (Collins *et al.*, 2008). לדוגמא, באמצעות שילוב טכנולוגיות ריצוף ו-Solexa הצליחו להראות שהבקטריה *Sulcia muelleri* עברה במהלך האבולוציה שינויים גנומים בדומה לחרק המאכסן איתו היא נמצאת בסימביוזה (McCutcheon and Moran, 2007). קולינס וחוב' (2008) השתמשו בטכנולוגיה זאת לאנליזת הטרנסקריפטום של הצמח הפוליפלואיד *Pachycladon enysii* (הורה רחוק שלו הוא allopolyploid ולכן מצופה שבגנום זה יהיו גנים משוכפלים). ע"י שילוב של מיפוי ו"כלי הרכבה" (assembly tools) הם הצליחו לקבוע את השתייכותם של הגנים המוכפלים לשני העתקי הגנום (Collins *et al.*, 2008). ריצוף בשיטת ה-mRNA-seq של Solexa שימש לאבחון דפוס ביטוי של גנים בשמר *S. pombe*. נמצאה שונות ביעילות ה-splicing בין אינטרונים שונים. בנוסף נמצא שעליה בשעתוק מובילה לעליה ב-splicing בגנים מסוימים (Wilhelm *et al.*, 2008). טכנולוגיית ריצוף של Solexa שימשה באורז לבחינת שינויים אפיגנטיים ברקמות של חייצים (tillers) ושוורשים. כן נמצאו miRNA ו-siRNA ייחודיים ברקמות שונות. יתרה מזאת, נמצא שירידה ברמת האנזים *mop1* (ההומולוג שלו בארבידופסיס הוא RNA-dependent RNA polymerase 2) גורמת לירידה תלויית רקמה של ה-miRNA ו-siRNA (Wang *et al.*, 2009).

## 2. מטרת המחקר

מטרת עבודה זאת הוא אפיון תהליכי יצירת השונות הסומקלונלית על ידי זיהוי הבדלים בין עצי תמר חריגים שמקורם בתרבית רקמה לעצים נורמאליים.

בעבודה זו התמקדתי בשני כיוונים:

2.1 אפיון דגם המתילציה ב-DNA תוך שימוש בשיטת ה-Bisulfite sequencing:

א. בדיקת דגם המתילציה במספר מקטעי DNA מהזנים 'ברהי' ו'מג'הול' אשר נמצאו בהם הבדלים בדגם

הפסים בין WT ל-Off Type בשיטת Methylation sensitive AFLP.

ב. בדיקת דגם המתילציה בגן DEFICIENS (DEF) שבודד מתמר נורמאלי בהשוואה לאותו גן

בפנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות'.

2.2 בחינת דגם הביטוי (הטרנסקריפטום) ב"לבבות" תמר על ידי ריצוף בהספק גבוה של ה-mRNA

השוואתו בין עצים נורמאליים מהזן 'מג'הול' לעצים בעלי פנוטיפ ננסי.

### 3. שיטות וחומרים

הפרק מחולק לשני תתי נושאים עיקריים: אפיון דפוסי המתילציה במקטעי DNA גנומים באמצעות שיטת ה-Bisulfite sequencing ובחינת דגם הביטוי בלבבות תמר באמצעות ריצוף בהספק גבוה של mRNA.

#### 3.1.1.1. מתילציה של DNA

##### 3.1.1.1. החומר הצמחי

דגמי המתילציה נבחנו בעצים מהזן 'ברהי' שסבלו מבעיות 'חנטה לקויה' ו'ריבוי שחלות' (כהן וחוב, 2003; Cohen *et al.*, 2004) ובמספר עצים ננסיים, מהזן 'מג'הול'. הדגם נבחן בעלים ופרחים שנשמרו בהקפאה עד השימוש. טבלה 2 מתארת את העצים ששימשו לאפיון דגם המתילציה.

טבלה 2 - עצים ששימשו לאפיון דגם המתילציה בשיטת ה-Bisulfite sequencing

שם הדוגמא	זן	מקור ריבוי	רקמה	פנוטיפ
4WT	ברהי	חוטר	עלים	נורמאלי
F1WT	ברהי	חוטר 92	תפרחת	נורמאלי
F2WT	ברהי	חוטר 92	תפרחת	נורמאלי
F4WT	ברהי	חוטר 98	תפרחת	נורמאלי
2	ברהי	תרבית רקמה	עלים	'חנטה לקויה' ו'ריבוי שחלות'
F5	ברהי	תרבית רקמה	תפרחת	'חנטה לקויה' ו'ריבוי שחלות'
F6	ברהי	תרבית רקמה	תפרחת	'חנטה לקויה' ו'ריבוי שחלות'
F7	ברהי	תרבית רקמה	תפרחת	'חנטה לקויה' ו'ריבוי שחלות'
21WT	מג'הול	חוטר	עלים	נורמאלי
22WT	מג'הול	חוטר	עלים	נורמאלי
24WT	מג'הול	חוטר	עלים	נורמאלי
13DW	מג'הול	תרבית רקמה	עלים	ננס
14DW	מג'הול	תרבית רקמה	עלים	ננס
18DW	מג'הול	תרבית רקמה	עלים	ננס

כל העצים ששימשו לאנליזה מקורם במטע יטבתה.

#### 3.1.1.2. הפקת DNA גנומי

DNA גנומי הופק מעלים ופרחים של תמר לפי Aitchitt וחובריו (1993) בשינויים כדלקמן: כשלושה גרם של רקמה נכתשו בחנקן נוזלי. החומר הכתוש הורחף ב-15 מ"ל בופר מיצוי (ראה הרכב הבופר בהמשך) והודגר משך 30 דקות ב-65 °C תוך ערבוב. המיצוי הראשון נעשה על ידי הוספת 15 מ"ל כלורופורם/איזואמיל-אלכוהול ביחס של 1:24 וסירכוז במשך 7 דקות ב-8,000 g. הפזה המימית הופרדה ועברה מיצוי נוסף על ידי הוספת נפח אחד של כלורופורם/איזואמיל-אלכוהול (1:24) וסירכוז באותם תנאים. לפאזה העליונה הוסף נפח אחד של איזופרופנול ונעשה סירכוז במשך 20 דקות ב-16,500 g. המשקע נשטף ב-20 מ"ל אתנול 70% וסורכוז משך 10 דקות ב-16,500 g. המשקע יובש, הורחף בשלושה מ"ל מים מזוקקים וטופל ב-1 µl RNase A משך 30 דקות ב-37 °C. חומצות הגרעין הושקעו ב-2.5 נפחים של אתנול (100%) לאחר הוספת Na-Acetate לריכוז סופי של 0.3 M וסירכוז למשך 10 דקות ב-16,500 g.

המשקע נשטף ב-20 מ"ל אתנול 70% וסורכז במשך 10 דקות ב-16,500 g. המשקע יובש והורחף ב-400  $\mu$ l מים מזוקקים. ריכוז ואיכות של ה-DNA הוערכו באמצעות הפרדה על ג'ל אגרוז 1% והשוואה עם כמות ידועה של סמן גודל (Lambda DNA / EcoRI+HindIII). איכות ה-DNA הוערכה על ידי יחס בליעת האור באורכי גל של 260 ו-280 ננומטר.

בופר מיצוי:

3% w/v CTAB (Hexadecyl Trimethyl-Ammonium Bromide)

1% v/v  $\beta$ -mercaptoethanol

1.4 M NaCl

20 mM EDTA (Ethylenediamine tetra-acetic acid)

100 mM Tris pH 8.0

### 3.1.3 הגברת מקטעים באמצעות PCR

הריאקציה בוצעה על תבנית של DNA גנומי ו/או cDNA שהופק מרקמות שונות של תמר מהזנים 'מג'הול' ו'ברה' עם תחלים שתוכננו מראש (טבלה 3) וסונתזו על ידי IDT (USA). הריאקציה בוצעה בנפח כולל של 25  $\mu$ l והכילה: 0.5  $\mu$ l dNTP (בריכוז של 10 pmol/ $\mu$ l), 10x PCR buffer 2.5  $\mu$ l (המכיל 15 mM  $MgCl_2$ ), 0.2  $\mu$ l Taq polymerase (בריכוז של 5 U/ $\mu$ l), 0.5  $\mu$ l מכל תחל (בריכוז של 20 pmol/ $\mu$ l), ו-10  $\mu$ l של DNA (בריכוז של 25 ng/ $\mu$ l) או cDNA (שהופק מ-1000 ng total RNA כמצוין בסעיף 3.2.2 ונמהל פי 10). נפח הריאקציה הושלם ל-25  $\mu$ l עם מים מזוקקים פעמים (ddH<sub>2</sub>O) מטופלים ב-DEPC. תנאי ריאקציית ה-PCR: חמש דקות ראשונות ב-95 °C, 40 מחזורים של: דקה אחת ב-94 °C, דקה אחת בטמפרטורת ההגברה (הטמפרטורה המתאימה ביותר לכל זוג תחלים) דקה אחת ב-72 °C ו-10 דקות אחרונות ב-72 °C. תוצרי PCR הופרדו על גבי ג'ל אגרוז בריכוז של 2%-3% כדי לוודא יצירה של תוצר יחיד.

### טבלה 3 - תחלים ששימשו לבידוד הגן DEFICIENS (DEF)

מספר	שם התחל	רצף התחל	אורך המקטע (bp)	מיקום המקטע ברצף הגן DEF
1	SPWs1F	GGCCTTAGTCTACTTTGAGAA	974	1-973
	SPWs1R	CCATATCCATAAGTAGAGATGATGC		
2	SPWs2F	GAGCAATTGAGTGTCTTAGAGATTAA	764	847-1611
	SPWs2R	CTACAAATCGTATATAAATTGCGTCG		
3	SPWs3F	GAGAGTGCCATCTAATAGAGTTGGTA	837	1479-2316
	SPWs3R	CTGCCGGTTAGTAGGATTTTCTATC		
4	SPWs6F	CGATTGACATTCTGTTTCCCAA	1029	3943-4972
	SPWs6R	CAAAGGTTTCCATGAAGAATGC		
5	SPWs7F	CTCATCAACCAAACATGTGTCC	1060	4843-5903
	SPWs7R	TGTTGAGTAGTGCGGTCAA		

מיקום המקטע ברצף הגן DEF	אורך המקטע (bp)	רצף התחל	שם התחל	מספר
5822-6456	634	GGGATAGATATCTCGACCATT	SPWs8F	6
		ATCTAAGCCCTTGCTACACA	SPWs8R	
6393-7279	886	TTGATCTCTTGCGGGTTTAAGA	SPWs9F	7
		TGGTATGAGGATGATGCTCCTAT	SPWs9R	
1770-2400	630	AACGAGAACCTCTAAAATCCACCGTGC	SPWs14F	8
		GATAAGCGAGACCTCGGCATCG	SPWs14R	
2100-2708	608	GCTTCTTTACCCTATCCTCCA	SPW3F	9
		ACAGTTTTTCCCCTTTTCTTTAG	SPW3R	
2281-2805	524	GATGGGGAGGGGGAAGATAG	SPW12F	10
		CGGAGGTTTCGGTTGATCTC	SPW12R	
3060-3743	683	GCATTAATACATGCCTGGTTACTGG	SPWs17F*	11
		CCACATCACATGATGAGAGCAG	SPWs17R*	
2182-3186	1004	CTTGCAACTTGAGTTGGTTCCA	SPWs18F*	12
		TTCCGCACCTTCCCAATGAT	SPWs18R*	
4145-4874	729	AGAACAGGTGATTCTGCGCTT	SPWs20F*	13
		CTTATTGGTCTTTTGCGGATGCGT	SPWs20R*	
1103-2251	1148	CTCTTGAGACTTCAAACAATTGAGG	SPWs21F*	14
		CTAACTTGGTAGCTGCACATGC	SPWs21R*	
29-1206	1177	TGCTTGATCCAAATTGCAGAACTG	SPWs22F*	15
		ACTTGAGTGAGTCACCTTAGGTAG	SPWs22R*	
	750	ATGGGGAGGGGGAAGATAGA	cSPWF**	16

F מייצג תחל forward, R מייצג תחל reverse. \* תחלים אלה שימשו לבידוד החלק האחרון (אקסונים 6 ו-7) של הגן DEF. \*\* תחל זה שימש לבידוד הגן מ-cDNA ותחל ה-reverse שלו היה SPWs17R.

### 3.1.4. בדיקת תוצרי ה-PCR

תוצרי ה-PCR הופרדו על גבי ג'ל אגרוז בריכוז של 2%-3% כדי לוודא יצירה של תוצר יחיד. ההרצה נעשתה במתקן אלקטרופורזה אופקי שהכיל בופר הרצה 1xTAE ובמתח של 100 V. כל דוגמא הכילה בופר הטענה (6x Loading Dye Solution), בנפח של 1/6 מכלל הדוגמא. GeneRuller של חברת Fermentas (U.S.A) שימש כסמן גודל. הג'לים צולמו במכשיר Fluorchem 8800 של חברת Alpha Innotech (U.S.A). אנליזת ריכוז התוצרים נעשתה באמצעות התוכנה AlphaEase FC. תוצרים בגודל הצפוי נחתכו מהג'ל ונוקו באמצעות Silica Bead DNA Gel Extraction Kit של Fermentas (U.S.A), בהתאם להוראות היצרן. התוצרים רוצפו והרצף אומת בהשוואה לרצף הצפוי.

### 3.1.5. קביעת רצף תוצרי ה-PCR

הריצוף נעשה על ידי מעבדות חי בע"מ ישראל ו/או על ידי חברת Macrogen בקוריאה. אנליזת הרצף ובדיקת ההומולוגיה נעשו בעזרת תוכנת BioEdit.

### 3.1.6. אפיון מתילציה במקטעי DNA בשיטת ה-Bisulfite sequencing

אפיון דגם המתילציה נעשה על רקמת עלים ותפרחות בשיטת ה-Bisulfite sequencing בעזרת ערכת (kit) EZ DNA Methylation Direct של חברת Zymo Research (U.S.A). השיטה מבוססת על טיפול ב-bisulfite שגורם לכך שכל C לא ממותל הופך ל-U ואילו C ממותל (5-methylcytosine) מוגן מהריאגנט ואינו משתנה. לאחר הטיפול, הדוגמא מוגברת ב-PCR וכל U הופך ל-T. ואילו 5-methylcytosine נשארים כ-C. יש להדגיש שלאחר ריאקציית הביסולפיט, שני גדילי ה-DNA כבר אינם משלימים זה לזה (Frommer *et al.*, 1992).

הרצפים נבחנו בתוכנת Methyl Primer Express v1.0 לזיהוי אתרים פוטנציאליים למיתול (אתרי CG, CNG ואיי מתילציה). לאתרים אלה נבנו תחלים ספציפיים אשר בדרך כלל לא הכילו אתרי CG ו-CNG. במקרים בהם לא נמצאו רצפים מתאימים לעיצוב התחלים, נעשה שימוש בתערובות של C ו-T (degenerate primers). לאחר ריאקציית ההיפוך ב-DNA בוצעה ריאקציית PCR עם תחלים מתאימים עבור כל מקטע (טבלה 4). תוצרי ה-PCR הורצו בג'ל אגרוז 2% ונוקו ממנו. ריכוז התוצרים המנוקים חושב על סמך המופע בג'ל.

הריאקציה בוצעה לפי הוראות היצרן: למבחנה שמכילה 500 ng DNA בתוך 20 µl הוסף 130 µl של CT Conversion Reagent. הדוגמאות הודגרו ב-98 °C למשך 10 דקות ולאחר מכן ב-64 °C למשך 2.5 שעות. הקולונות נשטפו ב-M-Binding Buffer 600 µl והוטענו בתוצרי ריאקציות המודיפיקציה. הקולונות נשטפו ב-M-Wash Buffer 100 µl, טופלו ב-M-Desulphonation Buffer 200 µl והודגרו בטמפ' החדר במשך 20 דקות. לאחר שתי שטיפות ב-200 µl של M-Wash Buffer, ה-DNA עבר מיצוי מהקולונה ב-M-Elution Buffer 10 µl וסרכוז 30 שניות ב-10,000 g.

#### טבלה 4 - תחלים ששימשו ל-Bisulfite sequencing

שם התחל	רצף התחל	אורך האמפליקון (bp)	טמפ' החיבור
תחלים לבחינת דגם המתילציה בגן DEF			
SPW_BS1F	GAAATGTGAAGTATGGAAAAGTGA	548	56
SPW_BS1R	CATCATTATTACAAATCCATAACAAA		
SPW_BS2F	GGATTATGAAGAAGGTTAAGGAG	499	56
SPW_BS2R	AAATTACTCCTCCACCTTATTC		
SPW_BS3F	TGAAGGATATAGGTATTAGTTA	567	56
SPW_BS3R	CTCCATAAATAATCTACTCTTC		
תחלים לבחינת דגם המתילציה בפרגמנטים מתמר שבודדו בשיטת MS-AFLP			
Bis1F	GYAGGGTTTAATTAATGTATGGA	123	56
Bis1R	ACAATAACTTTAAAATCATTCCA		
Bis2F	TTAATTTAATTTAAGAAAGGAAAAAG	276	56
Bis2R	CARAAATTTAATCTTTAAATC		
Bis5F	AGAYGGGTTAAAAGAGGGGT	236	56

שם התחל	רצף התחל	אורך האמפליקון (bp)	טמפ' החיבור
Bis5R	CCATTAACRACAAATATTAC		
Bis6F	GGTTATGAATTATAAAGTTTGAGAAT	542	56
Bis6R	CTTCCTCTTCTTTAACCTTTAAAT		
תחלים לבחינת דפוס המתילציה בגן FWA בארבידופסיס			
Met At FWA F	GATATTATTATGGAAGAAATAATTCGGAT	458	56
Met At FWA R	TTAAAAAGTAAAACGT TAAATATCTCRACACTTTTAACAATACCTT AAATAAATACAACAA		

F מייצג תחל Forward , R מייצג תחל Reverse.

בהפקת DNA גנומי מתקבלת תערובת מתאים רבים ומרקמות שונות. תאים אלה יכולים להיבדל בדפוס המתילציה שלהם. לכן, במהלך האנליזה של דפוסי המתילציה בשיטת Bisulfite sequencing, מקטעי ה-DNA הנבחרים עוברים הגברה, ומשובטים בחיידקים. עבור כל מקטע DNA רוצפו לפחות חמש מושבות מכל פרט נבחן. כל מושבה מייצגת את דגם המתילציה של מולקולת DNA אחת שהוגברה והוחדרה לחיידק בודד (Frommer *et al.*, 1992).

### 3.1.7. שיבוט התוצרים

כל מקטע עבר ליגציה לפלסמיד pGEM T easy של חברת Promega (U.S.A). השתמשנו ביחס מולארי של 1:3 בין המחדר לפלסמיד בהתאמה.

למבחנת הריאקציה הוספו:

5 µl 2x Rapid Ligation Buffer, T4 DNA Ligase

0.5 µl pGEM-T Easy Vector (50ng)

1 µl T4 DNA Ligase

השלמה עם מים עד נפח 10 µl

הריאקציה הודגרה ב-16 °C למשך הלילה.

כמות תוצר PCR שחושבה לפי הנוסחה:  $\frac{ng\ of\ vector\ x\ kb\ size\ of\ insert}{kb\ size\ of\ vector} \times vector\ molar\ ratio$

### 3.1.8. טרנספורמציה לחיידקים

הטרנספורמציה בוצעה לחיידקים קומפנטים מסוג DH5α של חברת Bio-Lab Ltd (Israel). הטרנספורמציה בוצעה לפי הוראות היצרן: חיידקי DH5α הופשרו 30 דקות על קרח. למבחנה עם כ-30 µl של חיידקים קומפנטים הוספו 5 µl מתוצר הליגציה למשך 30 דקות על קרח. המבחנות הועברו ל-42 °C למשך 90 שניות ולקרח למשך דקה וחצי. לאחר הוספת 1 ml של מצע גידול LB נוזלי, החיידקים הודגרו



למשך שעה ב-37 °C ונזרעו על צלחות עם מצע LB מוצק שמכיל אמפיצילין בריכוז 100 mg/l, 100 µl IPTG (מתמיסה בריכוז של 100mM) ו-20 µl X-Gal (מתמיסה בריכוז של 50mg/ml). DNA פלסמידי הופק ממושבות לבנות בערכה GeneJET plasmid miniprep kit של חברת (U.S.A) Fermentas.

## 3.2. ריצוף בהספק גבוה

### 3.2.1. החומר הצמחי

האנליזות הגנומיות נעשו על רקמת "לב" של תמר מזן 'מג'הול' של עצים נורמאליים שמקורם בחוטרים ועצים חריגים (שרובו בתרבית רקמה) בעלי פנוטיפ ננסי. שני עצים ננסיים בני שבע שנים ועץ נורמאלי בין ארבע נכרתו במטע נאות הכיכר בתאריך 23.01.2007 והועברו למעבדה. בנוסף נכרתו במטע יבניאל בתאריך 01.08.2007 עוד 14 עצים חריגים בני שבע שנים בעלי מופע ננסי, ובתאריך 22.03.2009 נכרת עץ נורמאלי נוסף (בן כ-25 שנה) ממתע מצפה שלם. במעבדה פורקו העצים עד למרכזם. "לב" התמר הוגדר כרקמה במשקל כ-100 גרם המצויה סביב המריסטמה הקודקודית. ה"לב" חולק לשני חלקים: "לב עליון" המכיל בעיקר את העלים העובריים הצעירים מאוד ואת המריסטמה הקודקודית, ו"לב תחתון" הנמצא מתחתיו, והמכיל את איזור ההתמיינות לצינורות ההובלה ולרקמת הפרנכימה של הגזע. שני החלקים נחתכו לחתיכות קטנות, הוקפאו בחנקן נוזלי ונשמרו ב-80 °C-. טבלה 5 מתארת את העצים הנורמאליים והעצים החריגים מזן 'מג'הול' בהם נבדק דגם ביטוי הגנים ברקמת ה"לב העליון".

טבלה 5 - עצי 'מג'הול' ששימשו לאנליזות הריצוף בנפח גבוה בשיטת Solexa

שם העץ	שיטת ריבוי	מטע	פנוטיפ
WT	חוטר	נאות הכיכר	נורמאלי
WT2	חוטר	מצפה שלם	נורמאלי
DW	תרבית רקמה	נאות הכיכר	ננס
DW1	תרבית רקמה	יבניאל	ננס
DW2	תרבית רקמה	יבניאל	ננס
DW3	תרבית רקמה	יבניאל	ננס
DW4	תרבית רקמה	יבניאל	ננס
DW5	תרבית רקמה	יבניאל	ננס
DW6	תרבית רקמה	יבניאל	ננס
DW7	תרבית רקמה	יבניאל	ננס
DW8	תרבית רקמה	יבניאל	ננס
DW9	תרבית רקמה	יבניאל	ננס
DW10	תרבית רקמה	יבניאל	ננס
DW11	תרבית רקמה	יבניאל	ננס
DW13	תרבית רקמה	יבניאל	ננס
DW14	תרבית רקמה	יבניאל	ננס
DW16	תרבית רקמה	יבניאל	ננס

הרקמה ששימשה לאנליזה הייתה "לב עליון".

### 3.2.2 הפקת Total RNA

הפקת Total RNA נעשתה בשיטת CTAB (Chang *et al.*, 1993) בארבעה מיצויים עם שינויים כדלקמן: כשלושה גרם של רקמה ("לב עליון"), נכתשו בחנקן נוזלי. החומר הכתוש הורחף ב-20 מ"ל בופר מיצוי (ראה הרכב הבופר בהמשך) והודגר משך 10 דקות ב-65 °C תוך ערבוב. המיצוי נעשה פעמים בנפח אחד של כלורופורם/איזואמיל-אלכוהול (1/24). הדוגמאות עורבבו במשך חמש דקות וסורכזו 10 דקות ב-14,000 g ב-15 °C. לפאזה העליונה הוסף 1/3 נפח 10 M LiCl והמבחנות עורבבו היטב והודגרו למשך הלילה ב-4 °C. הדוגמאות שוקעו במשך 20 דקות ב-14,000 g ב-4 °C. בשלב זה נעשו שני מיצויים נוספים: למשקע הוספו 1 מ"ל SSTE ו-1 מ"ל כלורופורם/איזואמיל-אלכוהול (1/24). הפאזה המימית הועברה למבחנות 2 מ"ל ונעשתה השקעה למשך חמש דקות ב-10,000 g. הפזה המימית מוצתה שנית בכלורופורם/איזואמיל-אלכוהול (1/24). ה-RNA הושקע לאחר חלוקת הפזה המימית לשתי מבחנות שבכל אחת כ-400 µl והוספת 133 µl של 10 M Amonium acetate ו-1,330 µl אתנול (100%). המבחנות עורבבו היטב והודגרו למשך לילה ב-20 °C. הדוגמאות סורכזו 15 דקות ב-16,000 g ב-4 °C. המשקע נשטף ב-1 מ"ל אתנול 70%, סורכזו 15 דקות ב-16,000 g, יובש והורחף ב-15 µl מים שעברו טיפול ב-DEPC (Diethylpyrocarbonate), שתי המבחנות של אותה דוגמא אוחדו. ריכוז ואיכות ה-RNA הוערכו באמצעות הפרדה על ג'ל אגרוז 2% ועל פי יחס בליעת האור באורכי גל של 260 ו-280 ננומטר.

הרכב התמיסות ששימשו בהפקה:

(1) בופר מיצוי:

0.1 M Tris-Cl pH 8.0

0.025 M EDTA pH 8.0

2 M NaCl

2% w/v CTAB

2% w/v Polyvinylpyrrolidone (PVP)

0.2% v/v β-Mercaptoethanol

(2) SSTE:

0.5% w/v SDS

1 M NaCl

0.01 M Tris-Cl pH 8.0

1 mM EDTA pH 8.0

### טיפול ב- DNase

ל-30 µg של Total RNA הוסף:

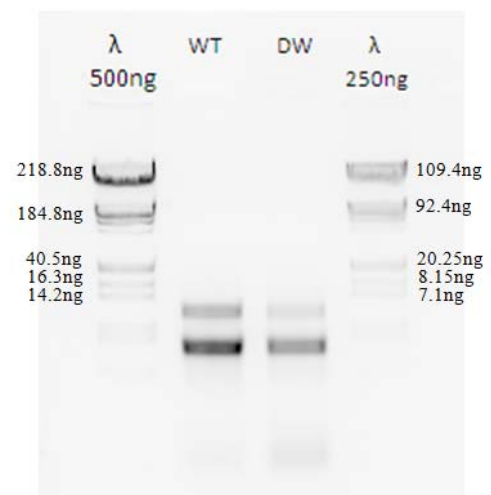
8 µl 10 x DNase Buffer

1 µl RNase inhibitor (40 U/ µl) (Fermentas, U.S.A)

1 µl RNase free DNase (1 U/ µl) (Promega, U.S.A)

נפח הריאקציה הושלם ל-40 µl ע"י מים שעברו טיפול ב-DEPC. הדוגמאות הודגרו במשך חצי שעה ב-37 °C. נפח הריאקציה הוגדל ל-500 µl עם מים שעברו טיפול ב-DEPC ונעשה מיצוי ב-250 µl פנול ו-250 µl כלורופורם-איזואמיל אלכוהול (1:24), המבחנות עורבבו וסורכזו 10 דקות ב-16,000 g ב-4 °C. לפאזה העליונה הוסף נפח שווה של איזופרופנול ו-5 µl גליקוגן (של חברת Fermentas) והדוגמאות סורכזו 15 דקות ב-16,000 g ב-4 °C. המשקע נשטף ב-1 ml אתנול 70% והושקע בסירכזו למשך 5 דקות ב-16,000 g ב-4 °C. הפלט הורחף ב-30 µl מים שעברו טיפול ב-DEPC.

### איור 1 - RNA מעץ תמר נורמאלי ונגס



רנ"א שהופק מריקמת "לב" של תמר מזן 'מג'הול' של עץ נורמאלי (WT) שמקורו בחוטר ומעץ חריג (DW) מתרבית רקמה בעל פנוטיפ ננסות. כסמן גודל שימש λ DNA התוך ב-*HindIII/EcoR*. יש לציין שסמן הגודל הוא מקטעי DNA בעוד שהמולקולות הנבדקות הן RNA.

### 3.2.3 יצירת cDNA

סינתזה של cDNA נעשתה בעזרת Verso™ cDNA Kit של חברת Thermo scientific (United Kingdom).

ל-1-5 µg של Total RNA הוספה 1 µl תערובת תחלים (פריימרים) random hexamers עד לריכוז סופי של 0.5-2 µM. התמיסה הודגרה 5 דקות ב-70 °C והועברה במהירות לקרח. למבחנת הריאקציה הוספו:

4 µl 5x cDNA synthesis buffer

2 µl dNTP Mix

1 µl RT Enhancer

1  $\mu$ l Verso Enzyme Mix

מים לנפח של 20  $\mu$ l

המבחנה הודגרה במשך שעה ב-42 °C. הריאקציה הסתיימה בהזגרה למשך 2 דקות ב-95 °C. הדוגמאות נשמרו ב-20 °C.

### 3.2.4 ריצוף בשיטת Solexa/Illumina

כל השלבים של אנליזת דגם ביטוי גנים בוצעו על ידי חברת FASTERIS בשווייץ. דגם הביטוי אופיין באמצעות שתי שיטות שונות. תהליך הריצוף זהה בשתייהן:

#### 3.2.4.1 Digital Gene Expression (DGE)

הדוגמאות הוכנו בעזרת ערכה (kit) "DGE-Tag Profiling Sample Prep Kit" של חברת Illumina תוך שימוש במספר הרכבים של אנזימי רסטריקציה. השיטה מבוססת על שימוש באנזימי רסטריקציה כך שמתקבלים תגים קצרים המייצגים לרוב את הקצה ה-3' של הגן.

##### 3.2.4.1.1 עיכול עם NlaIII

על גבי תבנית של 3  $\mu$ g mRNA סונתז גדיל ראשון של cDNA הקשור לחרוזים מגנטיים באמצעות אוליגו dT, בעזרת אנזימי SuperScript Reverse Transcriptase II. ה-mRNA עוכל על ידי RNase H ונבנה גדיל שני של cDNA על ידי אנזימי DNA Polymerase I. ה-cDNA הדו-גדילי נחתך באמצעות אנזימי הרסטריקציה NlaIII שחותך ברצף CATG. לאחר שטיפה בבופר נשארו רק המקטעים הקשורים לחרוזי אוליגו dT בקצה 3' (שלפני אתר החיתוך). לקצה החתוך של שני הגדילים חובר מתאם GEX NlaIII Adapter 1 הכולל בתוכו אתר רסטריקציה לאנזימי MmeI. אנזימי הרסטריקציה MmeI חתך לאחר 21 נוקליאוטידים מה-CATG, ובכך נוצר תג באורך 17 נוקליאוטידים (בנוסף לארבעה הבסיסים של אתר הרסטריקציה) שלא מחובר יותר לחרוזי אוליגו dT. לקצה החתוך חובר מתאם GEX Adapter 2, בעל רצף משלים לרצף שמחובר לתא (flow cell channels). מקטעי ה-cDNA חובקו משני צידיהם במתאמים (איור 2 א). ה-PCR נעשה עם תחלים מתאימים למתאמים. התוצרים הורצו על ג'ל, ונעשתה סלקציה למקטעים בגודל של 21 בסיסים.

##### 3.2.4.1.2 עיכול באנזימי רסטריקציה שונים

cDNA דו-גדילי עוכל בנפרד עם כל אחד מאנזימי רסטריקציה הבאים: *HinP1I*, *HpaII*, *HpyCH4IV* ו-*TaqI* (המשאירים קצוות דביקים 5' CG). לאחר שטיפה בבופר, תוצרי החיתוך מארבעת הריאקציות אוחדו וחוברו לחרוזים מגנטיים באמצעות אוליגו dT. לקצה החופשי חובר מתאם המכיל אתר רסטריקציה של *EcoP15I*. ה-cDNA עוכל עם האנזימי *EcoP15I* שחותך במרחק של 27 נוקליאוטידים מאתר ההכרה שלו. לאחר החיתוך השתחרר רצף באורך של 26-35 בסיסים. לקצה החתוך של התג (קצה 3') חובר מתאם שני

שריצופו משלים לרצף שמחובר ל-flow cell channels. התוצרים הופרדו על ג'ל לצורך בחירת מקטעים בגודל של 27 בסיסים או יותר. ניתן להבדיל בין ארבעת הריאקציות של אנזימי הרסטריקציה השונים על פי הבסיס הראשון שאחרי CG בתוך התג. בסיס זה משתנה בין ארבעת אנזימי הרסטריקציה.

#### **mRNA-seq 3.2.4.2**

הדוגמאות הוכנו בעזרת קיט "mRNA-seq" של חברת Illumina. השיטה מתבססת על שבירת ה-mRNA למקטעים וריצופם משני הכיוונים. מתקבל מידע על הטרנסקריפטום המלא. ה-mRNA הופרד מ-10 µg של Total RNA באמצעות poly-T מחובר לחרוזים מגנטים. ה-mRNA נשבר למקטעים קטנים בעזרת קטיונים ביולנטים בטמפרטורה גבוהה. סינתזת הגדיל הראשון של ה-cDNA נעשתה באמצעות האנזים Reverse Transcriptase ותחלים אקראיים. לאחר עיכול גדיל ה-RNA על ידי RNase H ויצירת גדיל שני של cDNA, נוצרו פרגמנטים של cDNA באורך של 400-50 בסיסים. יצירת overhang על קצוות ה-cDNA נעשה בעזרת T4 DNA polymerase ו-Klenow DNA polymerase. לאחר ליגציה של מתאמים ייחודיים לכל אחד מהקצוות, המקטעים הופרדו בג'ל ובודדו מקטעים באורך של 150-250 בסיסים.

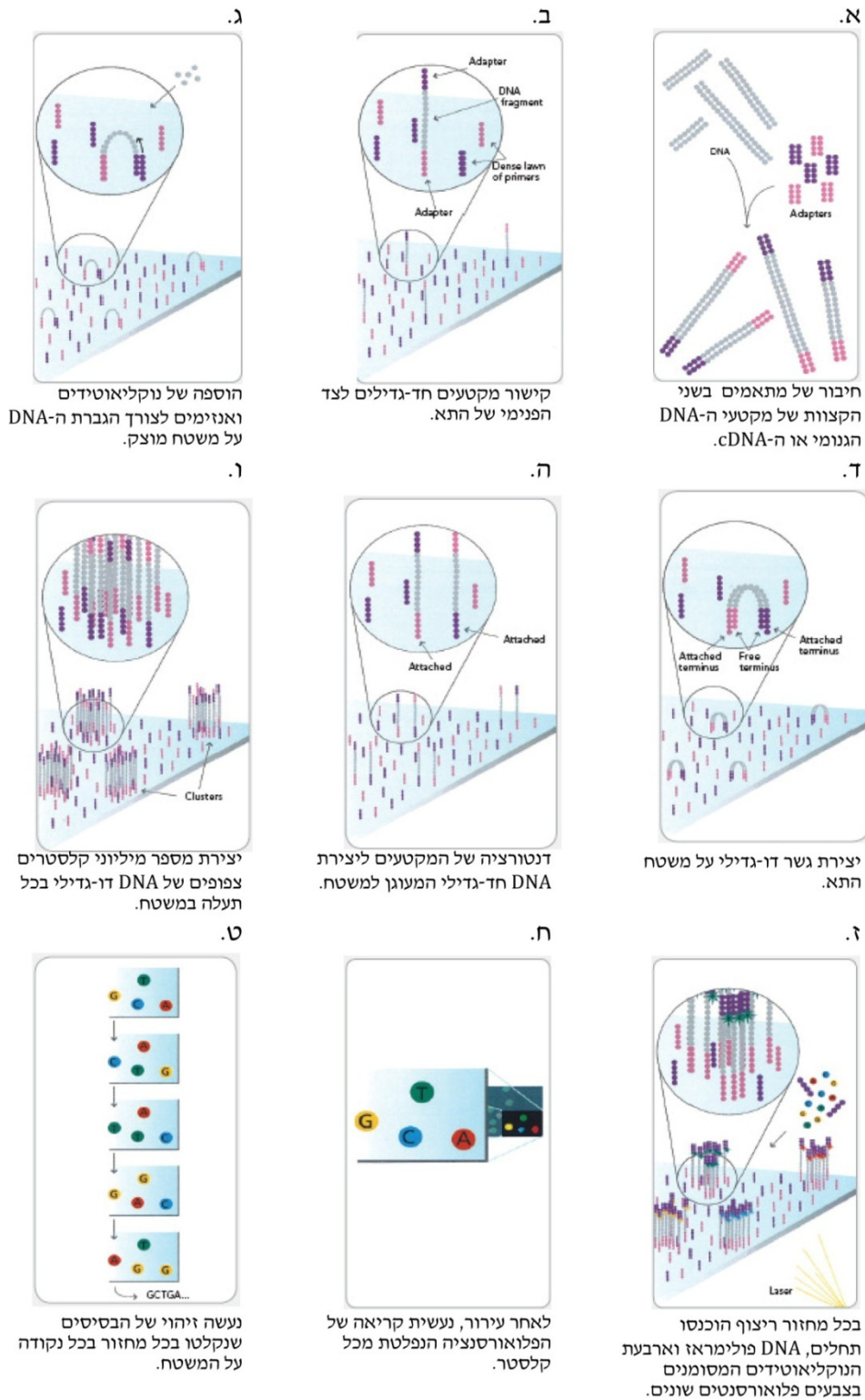
#### **תהליך הריצוף 3.2.4.3**

מקטעי ה-cDNA הופרדו לשני גדילים ונקשרו לצד הפנימי של תא ה-flow cell channels (התא מחולק למספר תעלות (channels) המאפשרות אנליזה של מספר דוגמאות במקביל) (איור 2 ב). המקטעים הוגברו על המשטח המוצק בעזרת נוקליאוטידים ואנזימים (איור 2 ג, ד). לאחר ההגברה נעשתה דנטורציה והתקבל DNA חד-גדילי מעוגן למשטח המוצק (איור 2 ה). נתקבלו כמה מיליוני קלסטרים צפופים ("מושבות DNA") של DNA דו-גדילי (איור 2 ו). בכל מחזור הגברה הוכנסו ארבעת הנוקליאוטידים, תחלים ו-DNA פולימראז. הנוקליאוטידים מסומנים בצבעים פלואורסנטים שונים והינם חסומים (Illumina משתמשת בנוקליאוטידים עם חסם רוורסיבילי (reversible terminators)). זהו נוקלאוטיד אנלוגי בעל מודיפיקציה המכסה בחלקה את הקבוצה 3'-OH שלו כדי לחסום זמנית את פעילות הפולימראז (בקה 3' שלהם) (איור 2 ז). בכל מחזור סינתזה נעשו ארבעה צילומים כדי לזהות את הבסיס שהוסף בכל מושבת DNA (איור 2 ח). שחרור החוסם נעשה בתגובה כימית או פוטוליזיס (Wu et al., 2007), הפלואורסנציה מוסרת ול-flow cell channels מוכנסת תערובת נוקליאוטידים חדשים למחזור ריצוף נוסף.

- לאחר סיום תהליך הריצוף, התמונות עובדו ב-Data Analysis Pipeline V. התהליך כולל 3 שלבים: בשלב הראשון הומרו התמונות לקבצי טקסט שבהם נרשמו העוצמות של כל אחד מארבעת הנוקליאוטידים בכל נקודה XY בתמונה. בשלב השני התבצע זיהוי של הבסיס בכל מחזור לפי נקודות XY (איור 2 ט). בשלב השלישי נוצר הרצף של כל מושבת DNA.

- ב-Digital Gene Expression ריצפו את כל התג שבאורך 21 או 27 בסיסים בהתאם לאנזים הרסטריקציה; ב-mRNA-SEQ ריצפו 36 בסיסים מכל כיוון של כל מקטע DNA (טכנולוגיה המכונה paired end), מהם 3 הבסיסים הראשונים היוו קוד לדוגמא (שהיה חלק מהאדפטורים).

איור 2 - טכנולוגיית הריצוף של Solexa/Illumina



התמונות לקוחות מהאתר של Illumina ([www.illumina.com](http://www.illumina.com)).

### 3.2.5. נירמול לתוצאות אנליזת ה-DGE

הנירמול בוצע ע"י ד"ר רון אופיר (מכון וולקני), באמצעות robust quantile normalization על הקריאות מהתגים של WT ושל Dwarf (Bolstad *et al.*, 2003). הנירמול מבוסס על ההנחה שברוב הגנים אין הבדלים משמעותיים בדגם הביטוי בין רקמה של עץ נורמאלי לשל עץ ננס, ולכן נקבעו החציון והשונות כקבועי ערך. בוצע מבחן שבודק את השערת האפס האומרת שהיחס בין הקריאות בכל תג זהות בין ה-WT ל-Dwarf. כיוון שדגם הביטוי נבדק במאות אלפי גנים (multi testing) נדרש תיקון למובהקות הסטטיסטית של הבדלים בין שני הפרטים. נעשו שני תיקונים: FDR ו-Bonferonni (Benjamini and Hochberg, 1995; Sokal and Rohlf, 1995). לצורך הצגת הנתונים (בדומה להצגה של תוצאות microarray), נעשתה טרנספורמציה לקבלת קירוב להתפלגות נורמאלית. הביטוי הדיפרנציאלי חושב לפי הנוסחה:  $\log_2(tD - nD + 0.5) - \log_2(nW + 0.5) - \log_2(tW - nW + 0.5) + \log_2(nD + 0.5)$  ו-tD הן הקריאות של תג ספציפי וסה"כ קריאות מהנתונים עבור Dwarf בהתאמה. ו-tW הן הקריאות של תג ספציפי וסה"כ קריאות מהנתונים עבור wt בהתאמה. תגים רבים בעלי ערכי טרנספורמציה נמוכים משש לא התפלגו בהתאם לעקומה הנורמאלית. מכיוון שבמערכת זו לא קיים מודל לטעות (error model) וכדי לוודא שהתגים שנבחרים באמת מייצגים הבדלים אמיתיים בביטוי, בוצע סינון (filtering) לתגים בעלי ערך טרנספורמציה גבוה משש. בנוסף נבחנו רק תגים בעלי ערכי p-value (הסיכוי לקבל תוצאה קיצונית מהתוצאה הנצפית, בהנחה שהשערת האפס נכונה) נמוכים 0.001 לאחר תיקון Bonferonni.

### 3.2.6. 5' RACE

האנליזה בוצעה ע"י חברת Fasteris בשווייץ באמצעות 5' RACE kit Invitrogen. שיטת ה-5' RACE (Rapid Amplification of cDNA Ends) משמשת להארכת מקטעי cDNA על ידי סינתזה באזור ה-5' של ה-mRNA. שיטה זאת דורשת ידע של רצף קצר ממקטע ה-cDNA הרלוונטי. לקצה 3' של cDNA נקשר הומופולימר dCTP. סינתזת הגדיל הראשון של cDNA על תבנית של רנ"א שליח נעשתה בעזרת תחל ספציפי GSP1 (טבלה 6), באוריינטציה הפוכה מהרצף שהתקבל באנליזת ה-DGE הגנומית של Solexa ותחל מתאים ל-poly C. אחרי הסינתזה של הגדיל הראשון עוכל הרנ"א השליח על ידי טיפול ב-RNase. נוקליאוטידים, GSP1 וחלבונים מיותרים שלא מחוברים ל-cDNA סולקו בעזרת קולונת S.N.A.P. תוצרי ה-5' RACE שובטו לווקטור TOPO והמקטעים רוצפו. תהליך זה מאפשר הארכה של רצף לא ידוע בין GSP1 לקצה 5'. התחלים ששימשו להארכה ב-5' RACE היו 100 תגים שונים שנבחרו מאנליזת ה-DGE (שיקולי הבחירה מוסברים בפרק 4.2.1.2).

## טבלה 6 - תגים שהוגברו ב-5' RACE

תחל - GSP1	שם הרצף	מספר
ACATCTTGCATCCTCCACA	ILA_seq1	1
CAATCATCACCCCTAGCACA	ILA_seq3	2
TCAGCAATACCAGGGAACA	ILA_seq5	3
TCCCTATACTCCTTGTACA	ILA_seq6	4
CTACCCATGTCAAGGACCA	ILA_seq7	5
GTGCTCTCCTGTTTCAGCCA	ILA_seq8	6
TCTACCATTACATGTAGCA	ILA_seq11	7
CCCAACCTCTACCATTACA	ILA_seq12	8
TCGAGCTAACTGTAGAACA	ILA_seq14	9
TCATCTTTGCACTATGGCA	ILA_seq17	10
CCAGTACTGGAAAACTCA	ILA_seq19	11
GCAAAATCAAATTACAACA	ILA_seq291	12
GTTGGATTTCTTTCATTCA	ILA_seq503	13
AAAAGTTCAATAGATACCA	ILA_seq510	14
CACCATTCTGAAGTAGGACA	ILA_seq517	15
ATTGAACTTTAGCCATTAATTCAATCG	ILA_CGaseq420	16
ATTCGATATTTGAAAATAATTTAAGCG	ILA_CGCseq438	17
TTTATAATCCATCCATCATCATATGCG	ILA_CGCseq466	18
TGCAGGAATCGCAGCTGCAGTTTGGCG	ILA_CGCseq495	19
ATTCATTTTCCACAGCTGCAACCAGCG	ILA_CGCseq499	20
ACTCACTTACAACCGTTCTCGGATCCG	ILA_CGGseq2	21
CACACACCCCAACCAAAATTCCAACCG	ILA_CGGseq134	22
CCGAACTTAGGAATAAACAGTATCCG	ILA_CGGseq135	23
CAAATTGCAACTGTGTAATGACACACG	ILA_CGTseq123	24

### 3.2.7. אנליזת BLAST

תוכנת Blast נמצאת באתר של NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). התוכנה משמשת לאיתור רצפים הומולוגים. במאי 2009 פורסמה טיוטה ראשונית של גנום התמר באתר <http://qatar.weill.cornell.edu/research/datepalmGenome/download.html>. מאוחר יותר הטיטה עודכנה, פורסמה במאגר המידע של NCBI (ACYX01234704 - ACYX01000001) וזוהו בה כ-19,414 גנים היפותטיים. ההבדל העיקרי בין שתי הטיטות הוא שהמקטעים מהטיטה הראשונה עברו הצלבה על ידי שימוש בריצוף paired-end ונוצרו scaffolds. הטענתי את טיוטת גנום דקל התמר בתוכנת BioEdit. החיפוש מול טיוטת הגנום של התמר נעשה "מקומית". נעשתה השוואה בין הרצפים בבסיס נתונים של nr/nt (Nucleotide collection) ו-Non-human non-mouse ESTs. כמו כן נעשה חיפוש ב-BlastX על בסיס תרגום רצף הבסיסים לרצף של חומצות אמינו.



### 3.2.8. אפיון דגם השעתוק באמצעות Q-RT-PCR

השיטה מאפשרת כימות של מולקולות RNA ומבוססת על SYBR Green - חומר פלואורסנטי שנצמד למולקולות DNA דו-גדיליות. במהלך PCR, SYBR Green נקשר ל-DNA דו-גדילי, וקומפלקס זה קולט אור כחול באורך גל של 480 nm ופולט אור באורך גל של 522 nm. האור הנפלט מיצג את כמות ה-DNA הדו-גדילי. מתוך כך ניתן לחשב את כמות ה-Template ששימשה בניסוי. ה-SYBR Green ששימש לריאקציה היה Absolute Blue QPCR SYBR Green ROX Mix של חברת Thermo scientific (United Kingdom). כימות התוצאות ב-Q-RT-PCR מסתמך על עקומת כיוול של הגן המנרמל והגן הנבדק המבוססת על סדרת מיהולים לריכוזים ידועים. עקומת הכיוול מבטאת יחס ליניארי בין  $C_t$  (מספר המחזורים המינימאלי של ריאקציית ה-PCR שבו הפלואורסנציה עוברת את ערך הסף) והכמות ההתחלתית של cDNA, ובכך ניתן לקבוע את כמות הגן הנבדק (Heid *et al.*, 1996).

עיצוב התחלים נעשה בתוכנות Primer Express ו-Primer 3 (טבלאות 7 ו-8). זוגות התחלים נבחרו כך שתוצר ה-PCR יהיה באורך של 80-120 בסיסים. טמפרטורת ה-Annealing הייתה  $59^{\circ}\text{C}$  לכל זוג תחלים שלא הכילו יותר מארבעה בסיסים זהים ברצף. אחוז ה-GC נע בין 40 ל-60. לכל רצף תוכננו 1-3 זוגות תחלים. התחלים נבדקו ב-PCR כדי לוודא קיום של תוצר יחיד. כל זוג תחלים שנתן תוצר יחיד בגודל המתאים והרצף שלו היה הומולוגי לרצף הצפוי, כויל לקביעת הריכוז האופטימאלי. כל תחל נמהל לריכוז של 2 pmol/ $\mu\text{l}$  עם מים "Ultra Pure Water". הריאקציה בוצעה בנפח כולל של 10  $\mu\text{l}$  והכילה תחלים (ראו לעיל), 5  $\mu\text{l}$  של Absolute blue SYBER Green ROX, 1  $\mu\text{l}$  של cDNA ו-Ultra Pure Water. בכל ריאקציות ה-Q-RT-PCR שימש ה-cDNA שנוצר מ-50 ng עד 4 pg של Total RNA כתבנית. ההגברה נעשתה במכשיר Rotor-Gene 6000 של חברת Corbett Life Science (Australia). נתאי ה-PCR היו: 15 דקות ב- $95^{\circ}\text{C}$ , 40 מחזורים של 10 שניות ב- $95^{\circ}\text{C}$ , 15 שניות ב- $59^{\circ}\text{C}$  ו-20 שניות ב- $72^{\circ}\text{C}$ . אנליזת התוצאות נעשתה בתוכנה Rotor-Gene 6000 Series Software 1.7.

#### טבלה 7 - תחלים ששימשו ל-Q-RT-PCR לבחינת תגים שאותרו ב-DGE

שם הרצף	שם התחל	רצף התחל	אורך המקטע
CGC438	CGC438s1F	GCAAGTTCCTTCAATTTACACGC	91bp
	CGC438s1R	GAAGATGGACACTCCTTTTAGCATG	
CGC438s2F	CGC438s2R	GCCCCAGAAGCAAGTTCCTT	113bp
		GCCAGGAAGATGGACACTCC	
CGT123	CGT123s1F	GACACACGTCCCAATTCAACAT	101bp
	CGT123s1R	CACGCCAGTTCCAACAATTA	
CGT123s2F	CGT123s2R	TGGCGTATGGGCTCTTTCTATT	101bp
		CGATAACCTCTTCCCCGTACCT	
Sol1	ILA1F	CGCAGCAAGACATTTTCTTCAC	81bp
	ILA1R	TTGAGAGGGAGACGGAAATTTG	
nSol1s2F	nSol1s2R	CACCGCAGCAAGACATTT	85bp
		CTTGAGAGGGAGACGGAAA	
Sol5	nSol5s3F	AATGGTGATGACCTGCCCAT	100bp
	nSol5s3R	TTATGTTGCCCTCGATTATGAGC	

שם הרצף	שם התחל	רצף התחל	אורך המקטע
Sol8	nSol8s1F	CACCGCAGCAAGACATTT	81bp
	nSol8s1R	AGAGGGAGACGGAAATTTGA	
	nSol8s3F	ACTTGCAGCCAGCTCCACA	110bp
	nSol8s2R	GAGGGAGACGGAAATTTGAGG	
Sol12	nSol12s1F	GACCGGGAAACAAGTTCAAT	112bp
	nSol12s1R2	AGAAGAAGGTGCGGAAAGTG	
Sol510	nSol510s1F	TGGACAGGACATCCATTCTAAG	101bp
	nSol510s1R	GAAGCTTTTGTCTCCCGTTT	
18s	18s3F	ACCACTGCGAAAGCATTTCG	100bp
	18s3R	CGTTTATGGTTGAGACTAGGACGG	

F מייצג תחל Forward , R מייצג תחל Reverse. בטבלה זאת מוצגים רק התחלים שנתנו תוצר יחיד ב-PCR.

### טבלה 8 - תחלים ששימשו ל-Q-RT-PCR לבחינת תגים שאותרו ב-mRNA-seq

שם הרצף	שם התחל	רצף התחל	אורך המקטע (bp)
ACYX01151934.1	gi261034334F	GGCACCTTCATGGACGTGTACT	104
	gi261034334R	GCAACAGCGTACGCAGTCTC	
ACYX01130212.1	gi261056360F1	AGAGGATTGCAGAGGAGTAGAGAC	104
	gi261056360R1	TTGAGTACAGGTCAGCATAGCTC	
ACYX01124357.1	gi261062215F2	GGACACGGAAAGGAAGAGAG	104
	gi261062215R2	TGGACGTGGAATACAAGAGGT	
ACYX01116467.1	gi261070105F4	CTATCAGGGTAACGCAACCA	105
	gi261070105R4	AAACCTGGAAGCCATTTGAG	
ACYX01112722.1	gi261073850F3	TGGTCAGCTCCTTACGCTATC	97
	gi261073850R3	CTTGATCTGCATGTTCGATCTTT	
ACYX01111891.1	gi261074681F	TCCAGACATCTGCCGGAATCTCA	100
	gi261074681R	GCAATTCACACTCTGATTCCCACC	
ACYX01110553.1	gi261076019F	ACGAAATGGACAGACAGCAG	98
	gi261076019R	GCTGTGAATCCTTTGGTTTG	
ACYX01107393.1	gi261079179F2	TTCAAGGCTGTTGGGCCACCT	104
	gi261079179R2	CCCTCCCTCTAAGGCTCATTTC	
ACYX01100129.1	gi261086443F2	TGGTTCGGCGGATTTGGTGAT	98
	gi261086443R2	TCTCACCTTGCCCGCTGATGTG	
ACYX01098460.1	gi261088526F	CCAAGGTGAACCTCCCTGAT	101
	gi261088526R	AAGAATCCCAGGTTGGTTTG	
ACYX01096809.1	gi261090177F	CCACGGCTCTTGCCAAAGC	99
	gi261090177R	AGCAGTTGCTGAGGGACTCCG	
ACYX01093611.1	gi261093375F	CCGTATAAGCTGACAACGGCGTA	104
	gi261093375R	ACTACATGCACAGGTGTGCC	
ACYX01091106.1	gi261095880F	ATTGCTGCAACTGATGCTGA	100
	gi261095880R	GTGCAACCTCAATCACAACG	
ACYX01090248.1	gi261096738F	AACAATGCCAACCGTGGTGC	101
	gi261096738R	CTAATCTCACCCCTGCGTTCCGG	
ACYX01087235.1	gi261099751F2	TGGGAGAGACAAGACTCCAAG	105
	gi261099751R2	TTCTGCTGCATTGTTTCATCA	

שם הרצף	שם התחל	רצף התחל	אורך המקטע (bp)
ACYX01078856.1	gi261108130F3	GTTAGATGGCCATTCGTTATCAT	104
	gi261108130R3	GATTGGTCAACCAAGAAACAGA	
ACYX01078076.1	gi261108910F2	AGGCTGTGTGGCACCTAACT	99
	gi261108910R2	AGTCCATGGCCATCTAAACG	
ACYX01065497.1	gi261121489F	AGAGCATTGAGCATAGGGCCATC A	101
	gi261121489R	TGGAAGTGGGCAACCAGTGT	
ACYX01061536.1	gi261125450F2	TGCTGCGAAGAACTTTGGTA	103
	gi261125450R2	AGGGAGTGCACATGTTGAAA	
ACYX01054383.1	gi261132603F	AAACAGTGTAACCCACAAAGAAA C	96
	gi261132603R	TCATAACTTTGGAATCCATGTAG G	
ACYX01049845.1	gi261137141F	TTTGGGATGCCTTTCGAGAA	98
	gi261137141R	ATGCGAACTGGCTGAGAGA	
ACYX01031616.1	gi261155370F	AAGCTAGTGCTGCTTTGAACC	100
	gi261155370R	TCACGACAGGAGGTTGAGTC	
ACYX01024895.1	gi261162286F	CTGAGAAGGAGTGGCAGTGGTTA	100
	gi261162286R	AGCCACTGCACACTCTCGAA	
ACYX01023841.1	gi261163340F2	GTTGAATGGTGCAGTGCTTC	103
	gi261163340R2	TTCTGTGAGCCCTGATTGTG	
ACYX01018104.1	gi261169077F2	TACTCGAGCTTGTTGACAGCCAC C	99
	gi261169077R2	TGCTTTACATCCTCAGCGCCCTC	

F מייצג תחל Forward , R מייצג תחל Reverse . בטבלה זאת מוצגים רק התחלים שנתנו תוצר יחיד ב-PCR.

כל זוג תחלים שנתן תוצר יחיד ב-PCR עבר כיוול שבמהלכו נבדק בריכוזים של 100 nM , 200 nM ו-300 nM כדי לקבוע את הריכוז האופטימאלי. הריכוז האופטימאלי נבחר כך שתהיה הגברה שלא תגביל את התקדמות הריאקציה. כמו כן נבדקה טמפרטורת ההתכה של כל תוצר. טמפרטורה זאת תלויה באורכו ובהרכב הבסיסים שלו. בדיקת טמפרטורת התכה מאפשרת לגלות אם יש זיהום, אם נוצר יותר מתוצר אחד, או אם קיימת קשירה לא ספציפית של התחלים. הריאקציה מתבצעת עם ביקורת שלילית ללא (NTC) Template. עקומת הכיול שנבנתה לכל זוג תחלים משמשת לכימות ריכוזו של גן. העקומה נבנית בריאקציות שנעשות עם מיהולים שונים של cDNA בשתים או שלוש חזרות. המיהולים היו או 1:1, 1:5, 1:25, 1:125, 1:625 או בחלק מהמקרים: 1:1, 1:4, 1:16, 1:64. גן ההשוואה היה 18S rRNA אשר נבחן לעקומת הכיול במיהולים הבאים של cDNA: 1:1000, 1:5000, 1:25000, 1:125000, 1:625000 או 1:4000, 1:16000, 1:256000, 1:64000. עקומת הכיול נבנתה מערכי ה- $C_t$  (threshold cycle) של כל מיהול כנגד ריכוז ה-Template בסקלה לוגריתמית. ערך  $C_t$  מוגדר כמספר המחזורים המינימאלי של ה-PCR שבו הפלואורסנציה עוברת את ערך הסף (threshold). ה-threshold נקבע בפאזה האקספוננציאלית של ה-PCR. ערך זה מצוי מעל רמת הרקע. השיפוע (M) של עקומת הכיול מעיד על הגברה אקספוננציאלית ואפקטיביות.

יעילות הריאקציה שווה ל:  $1 - 10^{-1/M}$ . הריאקציה אפקטיביות ב-100% כאשר בכל מחזור יש הכפלה של כמות התוצר. בתנאים אלה, ערך האפקטיביות הוא 1 וערך M הוא 3.3-.

חישוב כמויות יחסיות: הביטוי של הגן נקבע ביחס לגן הביקורת הריבוזומאלי 18S. החישובים נעשו בתוכנת Rotor-Gene 6000 Series Software 1.7 ו-Microsoft Excel. תחילה חושב היחס בין הריכוז ממוצע של הגן הנבדק לבין גן הביקורת. חישוב זה בוצע עבור דוגמאות עם cDNA של עץ נורמאלי ועץ ננס. החישוב נעשה על פי ממוצע הריכוזים במיהול של 1:5 או 1:4 המיהול של גן הביקורת היה 1:5000 או 1:16000. ערך Relative Expression חושב על ידי חלוקת הערך שהתקבל בעץ נורמאלי בערך שהתקבל בעץ הננס.

## 4. תוצאות

### 4.1. אפיון הבדלים בדגם מתילציה ה-DNA

#### 4.1.1. ביקורת חיובית לאפיון דגם המתילציה בשיטת Bisulfite sequencing

כביקורת חיובית לשיטת ה-Bisulfite sequencing בחנתי מתילציה בגן FWA מארבידופסיס (אקוטיפ *Columbia*). גן זה מכיל homeodomain ופקטור שעתוק שזוהה לראשונה בעקבות היכולת שלו לעכב פריחה. גן זה מבוקר באופן אפיגנטי. האקסון האחרון (אקסון מספר 8) בגן זה ידוע כבעל רמה גבוהה של מתילציה בשני אתרי CG (325) וחמש מתוך שבע ממותלות באתר CG (316), ראה איור 3. תוצאות אלה אישרו את אמינות הקיט ששימש לבחינת דגם המתילציה.

איור 3 - אתרי מתילציה שמורים בגן FWA בארבידופסיס

	290	300	310	320	330	340	350
gDNA sequence	AGGTGAAAATT	CAGATTCTGT	GAAGTTTCTT	CCTTGGGATTTT	CGATAGTGCC	CAGATGGAG	TAAATGGG
Methylated sequence	AGGTGAAAATT	CAGATTCTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	CAGATGGAG	TAAATGGG
Bis Arb 1	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 2	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 3	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 4	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 5	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 6	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	CGGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG
Bis Arb 7	AGGTGAAAATT	TAGATTTTGT	GAAGTTTTTTTT	GGGATTTT	CGATAGTGT	TAGATGGAG	TAAATGGG

קטע מאקסון מספר 8 בגן FWA. Bis Arb 1 עד Bis Arb 7 הם רצפים של שבע מושבות שנבדקו. האתרים CG316 ו-CG325 ממותלים בשיעור גבוה (Zhang *et al.*, 2006). דגם המתילציה נבחן בעלים של צמח ארבידופסיס אחד. לאחר ריאקציה ה-bisulfite הפרגמנטים שובטו לחידקים. מוצגים רצפים של שבע מושבות שנבחרו באקראי. gDNA Sequence מייצג את מקטע ה-DNA שעליו התבצעה הבדיקה. ה-Methylated sequence הוא הרצף הצפוי לאחר טיפול ב-bisulfite בהנחה שכל האתרים CG ו-CNG ממותלים (ולכן מוגנים ממודיפיקציה). צבע ורוד מייצג אתרי CG שנמצאו ממותלים.

#### 4.1.2. אפיון מתילציה של מקטעי DNA בשיטת ה-MS-AFLP

בבחינת דגם מתילציה ה-DNA ברקמת עלים באנליזת MS-AFLP שנעשתה על ידי ס. גורביץ (גורביץ, 2004, עבודת גמר) נמצא דגם פסים שונה בין עצים נורמאליים וחריגים (טבלה 1). על אף מספר ניסיונות לריצוף מקטעים אלה, רק בחמישה מתוך תשעה מקטעים נתקבלה איכות טובה של ריצוף. נעשו מספר ריצופים של חמשת המקטעים הנ"ל. לרצפים שהתקבלו מכל מקטע עשיתי ClustalW Multiple alignment בתוכנת BioEdit (כדי לאתר את המקטע ההומולוגי בין הרצפים הנבחרים ולקבל את הרצף המלא הארוך ביותר). בשלושה מקטעים (Bis1, Bis2, Bis6) התקבלו בכל אחד שני רצפים שונים לאחר שיבוט לפלסמיד וריצוף. כל הרצפים של כל אחד משני המקטעים האחרים (Bis5 ומקטע 7) היו הומולוגים ביניהם. בכל המקטעים המרוצפים נמצאו אתרי החיתוך של *MspI* ו-*EcoRI* מחוברים לאדפטורים שבאמצעותם נעשתה ריאקציה ה-MS-AFLP. לרצף שבין שני האדפטורים ביצעתי אנליזת Blast ב-BioEdit מול טיוטה ראשונית של רצף גנום התמר. בשלושה מחמשת המקטעים (Bis1, Bis2, Bis5) נמצאה הומולוגיה גבוהה לשני Scaffolds (מקטעי DNA) בטיוטת גנום התמר. אולם רק באחד

מה-Scaffolds ניתן היה ליצור תחלים משני צידי אתרי החיתוך של *MspI*. למקטע הרביעי (Bis6) נמצא הומוגיה גבוהה ב-Scaffold אחד. במקטע החמישי (מקטע 7, טבלה 1) לא נמצא אתר *MspI* בטיטת בגנום התמר (לפי דגם המתילציה שלו באנליזת MS-AFLP צפוי מקטע זה להיחתך באופן שונה בין העצים הנורמאליים והחריגים) ולכן לא המשכתי לבדוק רצף זה.

לארבעת המקטעים הנ"ל נבנו תחלים מתאימים שיכילו את האתר *MspI*. הם הוגברו ובודדו מ-DNA של עלים מעץ נורמאלי (WT) ועלים מעץ חריג ורופו. הרצפים הושוו ביניהם ולא נמצא כל הבדל בין ה-WT לעץ החריג ברצף הבסיסים.

דגם המתילציה במקטעים Bis1 ו-Bis2 נבדק בשלושה עצי 'ברה' נורמאליים ושלושה חריגים בעלי מופע של 'חנטה לקויה' ו'ריבוי שחלות' (טבלה 2). ב-Bis1 דגם המתילציה נבדק בתפרחות בלבד וב-Bis2 דגם המתילציה נבדק הן בתפרחות והן בעלים. דגם המתילציה במקטעים Bis5 ו-Bis6 נבדק בעלים של שלושה עצי 'מג'הול' נורמאליים ושלושה עצים ננסים (טבלה 2). דגמי המתילציה של מקטעים אלה מתוארים בטבלה 9.

בכל מקטע נבדקה המתילציה באתר החיתוך של *MspI* (CCGG) בו נמצא הבדל בדגם המתילציה בין עצים נורמאליים לחריגים (גורביץ, 2004, עבודת גמר) ובנוסף נבדק דגם המתילציה ביתר אתרי CG ו-CNG שנמצאו במקטע.

מקטע Bis1: אורך המקטע היה 123 בסיסים והוא הכיל שמונה אתרי מתילציה CG (הן ב-WT והן ב-Off-type). שישה אתרים מתוכם היו ממותלים בתדירות גבוהה ושני אתרים לא היו ממותלים כלל. באתר *MspI* לא נמצא הבדל ברור בין WT ל-Off-type ואחוז המתילציה היה דומה בשניהם (טבלה 9). בשלוש מושבות מעצי Off-type הייתה רמה נמוכה של מתילציה גם באתרי CNG. תוצאות אלה אינן מתאימות לתוצאות אנליזת ה-MS-AFLP, על פיהן האתר *MspI* במקטע Bis1 היה אמור להיות ממותל ב-WT ולא ב-Off-type (טבלה 1).

מקטע Bis2: מקטע באורך של 277 בסיסים המכיל 16 אתרי CG וארבע אתרי CNG. לא נמצאה מתילציה באתר *MspI* באף אחד מהפרטים. נמצאה מתילציה באתר CG אחד במושבה בודדת מפרט WT אחד בתפרחת. יתר אתרי ה-CG ו-CNG לאורך המקטע לא היו ממותלים (טבלה 9). תוצאות אלה אינן תואמות את אנליזת ה-MS-AFLP, לפיה הייתה המי-מתילציה (hemi methylation) באתר *MspI* בעלים של 'ברה' WT.

מקטע Bis5: מקטע באורך 236 בסיסים המכיל שבעה אתרי CG וחמישה אתרי CNG. שישה מתוך שבעה אתרי CG היו ממותלים בחלקם (בשני הפנוטיפים אך ברמה גבוהה יותר ב-WT) וכל אתרי ה-CNG לא היו ממותלים כלל. רק במושבה אחת מעץ ננסי הייתה מתילציה באתר *MspI*. תוצאות אלה אינן תואמות את אנליזת ה-MS-AFLP, לפיה אתר זה היה אמור להיות hemi-methylated בעצי DW ולא להיות ממותל בעצי WT. לעומת זאת, נמצא הבדל במתילציה בין WT ל-DW באתר 215. כל המושבות מעצי WT היו ממותלות באתר זה ורובן בעוד אתר אחד נוסף (204 או 187 או 198) (טבלה 9, איור 4). לעומת זאת רק ארבע מושבות (מתוך 15) היו ממותלות בעצי DW באזור זה. ממוצע האתרים הממותלים לאורך כל המקטע בעצי ה-WT היה 29.5% לעומת 5.7% בעצי DW (טבלה 9).

מקטע Bis6: מקטע באורך 542 בסיסים המכיל חמישה אתרי CG ו-18 אתרי CNG. אתר *MspI* (בבסיס 458) היה ממותל ב-95.6% ב-WT ורק ב-46.7% ב-DW (טבלה 9, איור 5). על פי אנליזת ה-MS-AFLP הייתה באתר זה hemi-methylation בעצי WT אבל לא בעצים הננסים (טבלה 1). באתר CG בבסיס 461 אחוז המתילציה ב-WT היה 100% ואילו ב-DW 46.7%. באתר CG בבסיס 454 אחוז המתילציה ב-WT היה 95.6% ואילו ב-DW 40%. באתר CG בבסיס 254 אחוז המתילציה ב-WT היה 88.9% ואילו ב-DW 40%. אתר 102 היה ממותל בכל עצי ה-WT וה-DW ב-100%. אתר CNG בבסיס 84 היה ממותל במושבה אחת של פרט DW. סה"כ אחוז מתילציה באתרי CG במקטע זה ב-WT היה 96% ובעצי DW 54.7%. זה המקטע היחיד בו נמצאה התאמה בין תוצאות המתילציה משתי האנליזות.

### טבלה 9 - דגמי מתילציה של ארבעה מקטעי ה-DNA שאופיינו באנליזת MS-AFLP

#### א. מקטע Bis1 ('ברהי')

אורך מקטע (bp)	אחוז אתרי CG ממותלים במקטע כולו	אחוז מתילציה באתרי CG ספציפיים						מספר * אתרי CG ממותלים	פנוטיפ
		95	92 ( <i>MspI</i> )	88	77	74	54		
123	25.8	56.7	51.1	12.2	31.1	61.1	0	(8) 6	WT
	30.4	37.2	39.4	26.1	31.7	80.0	29		Off-type

#### ב. מקטע Bis2 ('ברהי')

אורך מקטע (bp)	אחוז CG ממותל	אחוז מתילציה באתרי CG ספציפיים		מספר * אתרי CG ממותלים	פנוטיפ
		( <i>MspI</i> ) 115	251		
277	0.4	0	6.7	(16) 1	WT
	0	0	0		Off-type

#### ג. מקטע Bis5 (מג'הול)

אורך מקטע (bp)	אחוז CG ממותל	אחוז מתילציה באתרי CG ספציפיים						מספר * אתרי CG ממותלים	פנוטיפ
		215	204	198	187	63	( <i>MspI</i> ) 45		
236	29.5	100	75.6	12.2	18.9	0	0	(7) 6	WT
	5.7	26.7	0	0	0	13.3	7		Off-type

#### ד. מקטע Bis6 (מג'הול)

אורך מקטע (bp)	אחוז CG ממותל	אחוז CG מתילציה באתרים					מספר * אתרי CG ממותלים	פנוטיפ
		461	458 ( <i>MspI</i> )	454	254	102		
542	96	100	95.6	95.6	88.9	100	(5) 5	WT
	54.7	46.7	46.7	40	40	100		Off-type

התוצאות מבוססות על ממוצע של שלושה פרטים מכל פנוטיפ. בחינת דגם המתילציה נעשה על רקמת תפרחות במקטע Bis1, ועל רקמות עלים ותפרחות במקטע Bis2 ורק ברקמת עלים עבור המקטעים Bis5 ו-Bis6. בכל פרט רוצפו בין שתיים לשש מושבות. המספרים שבכותרת מציינים את מספר הבסיסים ברצף בו מתחיל האתר CG \* מוצגים רק אתרי CG בהם הייתה מתילציה לפחות במושבה אחת. בסוגריים מצוין מספר כלל אתרי ה-CG במקטע. אתר *MspI* הינו אתר רסטריקציה שבו נמצא דגם מתילציה שונה באנליזת MS-AFLP.

איור 4 - דגם המתילציה במקטע Bis5

	170	180	190	200	210	220	230
gDNA	CTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGCCGATTCCGTTATTGAGGCGGCAACACTTGCCGCTAATGG						
Bisulfite Modification	CTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGTCGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
ע							
מושבה							
22WT - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTCGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
22WT - 2	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
22WT - 3	TTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
22WT - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
22WT - 5	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
22WT - 6	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTCGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
21WT - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
21WT - 2	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
21WT - 3	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
21WT - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTCGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
21WT - 5	TTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
24WT - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
24WT - 2	TTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
24WT - 3	TTGGTGGTAATATGAGGGGAAGGGGTCGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
24WT - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
24WT - 5	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
18DW - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
18DW - 2	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
18DW - 3	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
18DW - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
18DW - 5	TTGGTGGTAATATGAGGGGAAGGGGTTGGGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
13DW - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
13DW - 2	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
13DW - 3	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
13DW - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
13DW - 5	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
14DW - 1	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
14DW - 2	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
14DW - 3	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
14DW - 4	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						
14DW - 5	TTGGTGGTAATATGAGGGGAAGGGGTTGAGTGAGGGTTGATTTTCGTTATTGAGGCGGTAATATTTGTCGTTAATGG						

gDNA מייצג קטע מהרצף בו התגלו מרבית האתרים הממותלים. הרצף היה זהה ב-WT ו-Off type Bisulfite Modification. מייצג את הרצף הצפוי בהנחה שכל אתרי CG ממותלים. צבע ורוד מייצג אתרי CG ממותלים וצבע זהוב מייצג אתרי CNG ממותלים.



איור 5 - דגם המתילציה במקטע Bis6

	90	100	250	260	450	460	470
gDNA- 1	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....	..... ..... ..... ..... ..... .....
Bisulfite Modification	ATGCTGACTTGGTACTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	CAGGCTTAATACCCGATACAC	CAGGTTAATATTCGATATAT	GTGTTCTGACGGCCGGCGGCATTTTTAGC	GTGTTCTGACGGTCGGCGGTATTTTTAGT	GTGTTCTGACGGTCGGCGGTATTTTTAGT
מושבה עץ							
22WT - 1	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
22WT - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
22WT - 3	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
21WT - 1	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
21WT - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
21WT - 3	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
21WT - 4	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
21WT - 5	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
24WT - 1	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
24WT - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
24WT - 3	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
24WT - 4	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
24WT - 5	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
<hr/>							
18DW - 1	ATGCTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
18DW - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
13DW - 1	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
13DW - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
13DW - 3	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
13DW - 4	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
13DW - 5	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT	GTGTTTTGACGGTCGGCGGTATTTTTAGT
14DW - 1	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
14DW - 2	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
14DW - 3	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
14DW - 4	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT
14DW - 5	ATGTTGATTTGGTATTTAAAGCGAATGT	ATGTTGATTTGGTATTTAAAGCGAATGT	TAGGTTAATATTCGATATAT	TAGGTTAATATTCGATATAT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT	GTGTTTTGATGGTTGGTGGTATTTTTAGT

מוציגים שלושה קטעים מהרצף בהם התגלו מרבית האתרים הממותלים. gDNA מייצג את הרצף הנבחן. הרצף היה זהה ב-WT ו-off type. Bisulfite Modification מייצג את הרצף הצפוי בהנחה שכל אתרי CG ממותלים. צבע ורוד מייצג אתרי CG ממותלים וצבע צהוב מייצג אתרי CNG ממותלים. צבע תכלת על Y מסמן שבריצוף התקבל בבסיס זה סיגנל כפול של C ו-T. צבע אפור מציג אתר *MspI* (CCGG) בו צפויים הבדלים במתילציה על פי MS-AFLP.

### 4.1.3. אפיון דגם המתילציה של הגן DEF DEFICIENS (DEF)

מאור שוחט מצאה באנליזה של Q-RT-PCR שרמת הביטוי של הגן DEF בתפרחת הייתה נמוכה פי שלושה בעצים חריגים בעלי פנוטיפ "חמור" של 'חנטה לקויה' ו'ריבוי שחלות' לעומת העצים הנורמאליים (שוחט, 2007, עבודת גמר). בבחינת דגם המתילציה ברקמת עלים על ידי אנזימי רסטריקציה *HpaII* ו-*MspI* מצאה מאור שוחט לעיתים תוצר חלש מאוד לאחר חיתוך ב-*MspI* בעצים חריגים בלבד. ממצא זה עשוי לנבוע ממתילציה ברמה נמוכה של אתר החיתוך (שוחט, 2007, עבודת גמר).

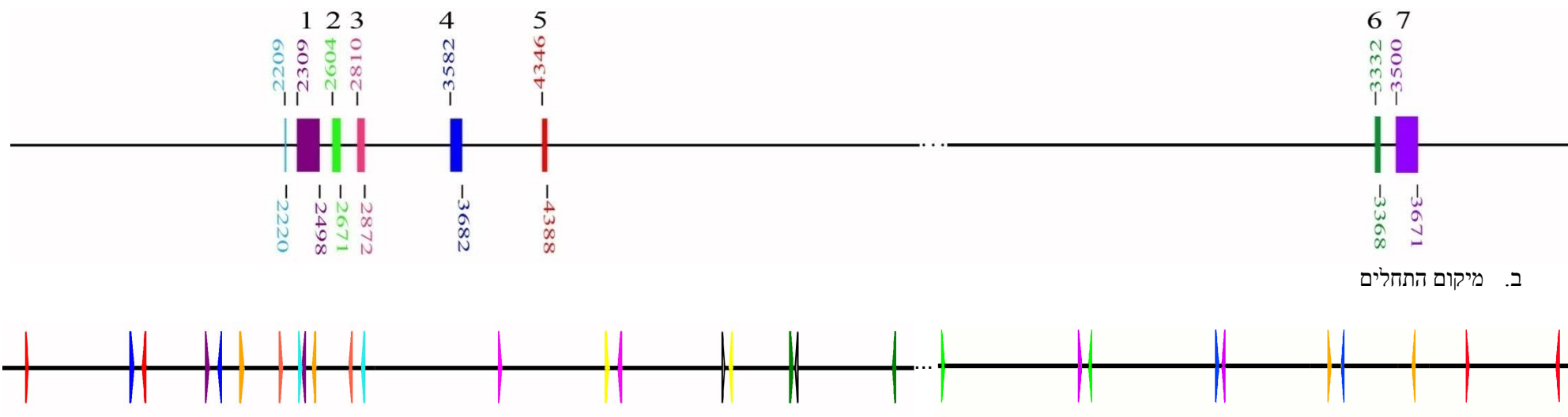
בעבודה זאת נבחן דגם המתילציה ברקמת עלים ותפרחות של עצים נורמאליים ובעלי פנוטיפ "חמור" של 'חנטה לקויה' ו'ריבוי שחלות' באזורים נרחבים של הגן DEF בשיטת ה-Bisulfite sequencing. השלב הראשון היה בידוד הגן מגנום התמר. לצורך כך איתרתי במאגר המידע של NCBI גנים אורתולוגים באורגניזמים חד פסיגים כמו אורז, תירס, אספרגוס דקל השמן וכן בארבידופסיס. באורז גן זה נקרא SPW1, בתירס Silky1, באספרגוס AODEF, בדקל השמן DEF ובארבידופסיס APETALA3. בכל האורגניזמים הנ"ל הגן בנוי משבעה אקסונים (איור 6). האקסון הראשון והאחרון הם הארוכים ביותר. באורז, אספרגוס וארבידופסיס אורך האקסון הראשון הוא 187 בסיסים, השני הוא 66, השלישי הוא 61, הרביעי הוא 99, החמישי הוא 41, השישי הוא 44 ואורך האקסון האחרון הוא 170, 173 ו-194 בסיסים בהתאמה. אורכי האינטרונים (חוץ מהאינטרון החמישי) אינם זהים ונעים בין 100 ל-450 בסיסים. אורך האינטרון החמישי באורז הוא 2168 בסיסים ובאספרגוס אורכו 4665 בסיסים. לעומתם, בארבידופסיס אורך אינטרון זה הוא 279 בסיסים. לצורך חיפוש אזורים שמורים בין האורגניזמים השתמשתי בתוכנות Blast ו-BioEdit. בין האורגניזמים החד פסיגיים - אורז, תירס, אספרגוס ודקל השמן נמצאה הומולוגיה גבוהה באקסונים ואילו ההומולוגיה לארבידופסיס הייתה ברמה הרבה יותר נמוכה (איור 7, טבלה 10). באינטרונים ובאזורי 3' UTR ו-5' UTR לא נמצאה הומולוגיה בין האורגניזמים שנבדקו. נעשה Blast בין הגן DEF מדקל השמן לטיוטת גנום התמר שפורסמה לאחרונה (ACYX01000001-ACYX01234704). נמצאו שלושה קונטיגים בעלי הומולוגיה משמעותית לגן. בקונטיג הראשון (ACYX01106400.1) נמצאה הומולוגיה לחמשת האקסונים הראשונים, בקונטיג השני (P.0S\_1319678) נמצאה הומולוגיה לחלק מהאקסון הראשון (קונטיג זה לא נמצא בטיוטת הגנום המעודכנת, אך נמצא בטיוטה הראשונית שפורסמה). הוא מכיל רק 107 בסיסים והומולוגי לחלק הראשון של האקסון הראשון. האזור ההומולוגי בין שני קונטיגים אלה הוא באורך של 13 בסיסים בלבד. בקונטיג השלישי (ACYX01099468.1) נמצאה הומולוגיה לאקסונים שש ושבע. לאחר הגברה של מקטעים מקונטיגים אלה וריצופם (ראה להלן), הוכח הקשר שבין הקונטיגים השונים לאזורים השונים בגן DEF. בעזרת האלגוריתם TSSP בתוכנה Softberry ([www.softberry.ru](http://www.softberry.ru)) נמצא אזור מועמד ל-TATA box שנמצא 99 בסיסים לפני קודון תחילת התרגום (start codon) ובעל רצף CTTCTCTAAAA. על בסיס הרצפים של גנום התמר תוכננו מספר זוגות תחלים (טבלה 3). ריאקציית PCR בוצעה על תבנית DNA גנומי של עלי 'בריה' נורמאליים וחריגים. סה"כ בודדו שני רצפים המורכבים ממקטעים קצרים יותר: אחד באורך 7074 בסיסים הכולל 2273 בסיסים לפני קודון התחלת התרגום, חמישה אקסונים ראשונים וחלק מאינטרון חמש. רצף שני באורך 4848 בסיסים שמכיל חלק מאינטרון חמש,

אקסונים שש ושבע ו-1203 בסיסים לאחר קודון הפסק. למרות מאמצים רבים לא הצלחנו לבודד כמה אזורים מה-DNA הגנומי של הגן וביניהם את מרבית האינטרון חמש, 473 בסיסים באזור ה-UTR 3', 1168 בסיסים המכילים את סוף אקסון שלוש (12 בסיסים אחרונים), אינטרון שלוש, אקסון ארבע וחלק מאינטרון ארבע. אזורים אלה עשירים מאוד בחזרות של הבסיס T. לעומת זאת הצלחתי לבודד את הרצף של האקסונים שלוש וארבע מ-cDNA שהופק מרקמת עלים מתמר נורמאלי.

בהשוואת הרצפים נמצאו כמה הבדלים ברצף בין טיוטת הגנום (של הזן 'חלאס') לבין הרצף הנבדק מהזן 'ברהי': אזור 1990-1945 (UTR 5' של הגן) הוא אזור בעל חזרות GA והוא ארוך יותר ב'ברהי' בהשוואה ל'חלאס'. בזן 'ברהי' החזרות הן באורך של 55 בסיסים ואילו ב'חלאס' הם באורך 20 בסיסים. בהשוואה בין הרצפים נמצא שבאתרים 1887 ו-1898 מצוי הבסיס A ברצף בזן 'חלאס' והבסיס C בזן 'ברהי'. באתר 2058 יש T בזן 'חלאס' ו-C ב'ברהי'. אתרים אלה נמצאים באזור ה-UTR 5'. בנוסף, ב'ברהי' נמצא רצף של 147 בסיסים שמתחיל ב-UTR 5' ונמשך לאקסון הראשון רצף זה לא מופיע ב'חלאס' (או כתוצאה מבעיות ריצוף או שהאקסון הראשון בו קצר יותר). בזן 'ברהי' קיים רצף של 50 בסיסים באינטרון ארבע שלא קיים ב'חלאס'. באתר 5268 באקסון החמישי ב'ברהי' יש זוג בסיסים CG שאינם קיימים בזן 'חלאס'. כל הרצפים הנ"ל היו זהים לחלוטין בעץ WT ובעץ בעל פנוטיפ של 'חנטה לקויה' וריבוי שחלות'.

## איור 6 - מבנה הגן DEFICIENS

א. מבנה הגן

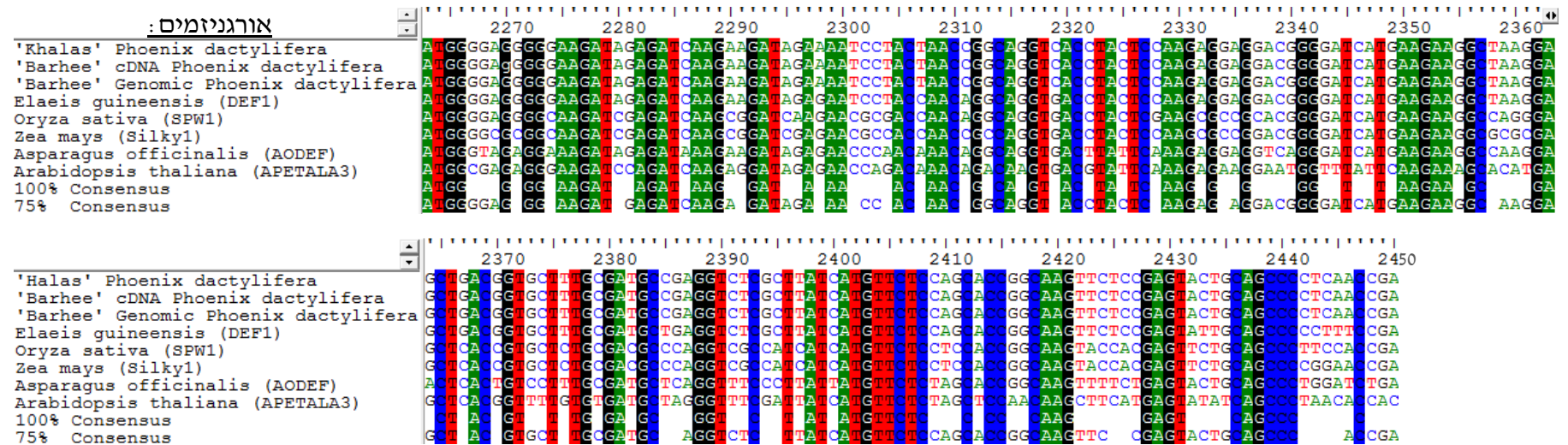


ב. מיקום התחלים

א. מוצגים שני מקטעים המכילים את מרבית הגן: מלבנים צבעוניים מסמנים את שבעת האקסונים של הגן. המספרים מציגים את מספרי הבסיסים (מעל לאקסון – התחלה, מתחת לאקסון – סיום). TATA Box מסומן בתכלת.

ב. החצים הצבעוניים מציינים את זוגות התחלים ששימשו לריצוף הגן. כל זוג תחלים מסומן בצבע ייחודי.

איור 7- רצף אקסון מספר 1 מהגן DEF באורגניזמים שונים



האיור מתאר הומולוגיה של האקסון הראשון בגן DEF בין DNA גנומי ו-cDNA של תמר מזן 'ברהל', DNA גנומי של תמר מזן 'חלאס' וכן הרצפים האורתולוגיים מדקל השמן, אורז, תירס, אספרגוס וארבידופסיס. בשתי השורות האחרונות מוצגים רצפי הקונסנזוס לפי 100%-ו 75% זהות בין כל הרצפים שנבחנו.

טבלה 10 – השוואת רצף ה-cDNA של הגן DEF בין התמר לאורגניזמים אחרים

חלבון		cDNA		אורגניזמים (שם הגן)
אחוז הומולוגיה	אורך מקטע (aa)	אחוז הומולוגיה	אורך מקטע (bp)	
90	225	91	678	<i>Elaeis guineensis</i> (DEF)
83	225	82	678	<i>Asparagus officinalis</i> (AODEF)
75	224	78	675	<i>Oryza sativa</i> (SPW1)
75	227	76	684	<i>Zea mays</i> (silky1)
54	232	65	699	<i>Arabidopsis thaliana</i> (AP3)

שיעור ההומולוגיה ברמת ה-cDNA וברמת החלבון של הגן DEF מתמר וגנים אורתולוגים במינים אחרים נקבע באמצעות Blast. שיעור ההומולוגיה ברמת ה-cDNA נקבע ב-BlastN. שיעור ההומולוגיה ברמת החלבון נקבע ב-BlastX. הגן הנבדק בתמר היה באורך 225 חומצות אמינו.

לצורך אפיון דגם המתילציה של מקטע מהגן DEF, נעשה חיפוש לאזורים המכילים איי מתילציה (אזור עשיר באתרים CG ו-CNG) בעזרת תוכנת Methyl Primer Express v1.0. החיפוש נעשה בשני הרצפים שבודדתי מתמר ואותרו האזורים הבאים:

1. מקטע באורך 7047 בסיסים מ-UTR 5' ועד סוף אקסון חמש (איור 8 א).
2. מקטע באורך 4848 בסיסים מאקסון שש עד ה-UTR 3' (איור 8 ב).

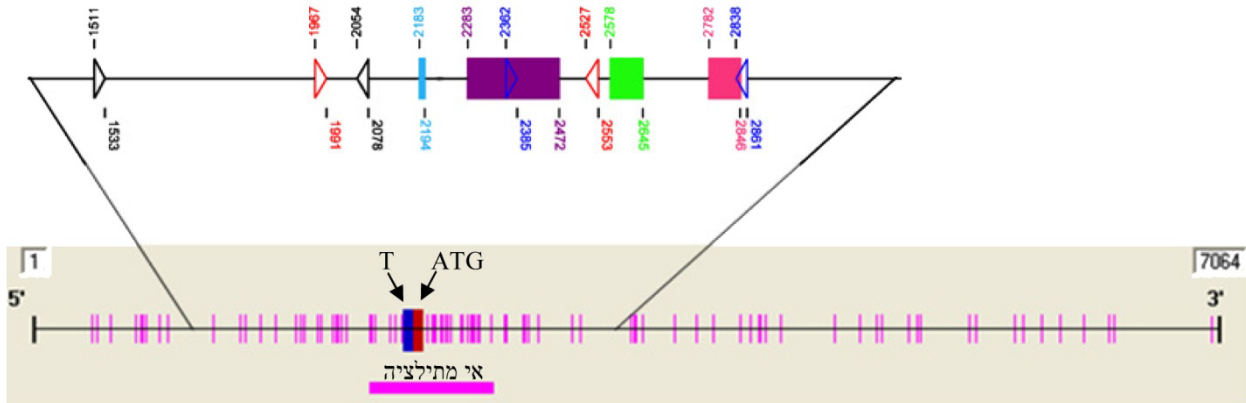
באזור הראשון נמצא "אי מתילציה" אחד באורך של 760 בסיסים שמתחיל בבסיס מספר 1996 ונגמר בבסיס 2756. נבנו שלושה זוגות תחלים (טבלה 4, איור 8 א) בעלי אזורים חופפים לצורך בידוד מקטע אחד באורך של 1350 בסיסים. מקטע זה כולל את ה-TATA box אזורים המועמדים להיות 5' UTR ואזורים מקודדים מהאקסון הראשון ועד האקסון השלישי. במקטע זה מצויים 41 אתרי CG ו-30 אתרי CNG (טבלה 11).

המתילציה נבחנה על רקמת עלים ותפרחות של שלושה עצי תמר נורמאליים ושלושה עצי תמר בעלי פנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות' מהזן 'ברהי'. בכל המושבות שרוצפו בשיטת Bisulfite sequencing מכל הצמחים נמצאו רק שלושה אתרי CG ושני אתרי CNG ממותלים. באתרי CG 1828 ו-2473 שמקורן בעצי WT נמצאה מתילציה בשתי מושבות מתוך 26 מושבות (לא נמצאה מתילציה באתרים אלה בעצי off-type). באתר CG 2404 נמצאה מתילציה במושבה אחת (מתוך 26) של עצי WT ובמושבה אחת (מתוך 24) של עצי off-type. בשתי מושבות משני עצי WT שונים נמצאה מתילציה באתר CNG 1992 ובמושבה אחת מעץ WT אחד נמצאה מתילציה באתר CNG 1983. כל המושבות בהן נמצאה מתילציה היו מרקמת תפרחות. לסיכום, במקטע הנבדק

שאוּרכו 1350 בסיסים לא נמצאה רמת מתילציה משמעותית ולא נמצא הבדל בדגם המתילציה בין עצי WT לעצי off-type (טבלה 11).

### איור 8 - אזורי מתילציה בגן DEF ותחלים שתוכננו לריאקצית Bisulfite sequencing

א. רצף DEF מ-UTR 5' ועד אקסון 5



- אקסון מספר 1, - אקסון מספר 2, - אקסון מספר 3, - איזור TATA box משוער.  
 - זוג תחלים SPW\_BS1, - זוג תחלים SPW\_BS2, - זוג תחלים SPW\_BS3.  
 קו וורוד מאונך מציין אתר CG. אי מתילציה מיצג אזור בו יש לפחות 300 אתרי CG בין הבסיסים 1996 ל-2737. ATG מציין את קודון תחילת התרגום, T מציין את ה-TATA box.

ב. רצף DEF מהאקסון 6 עד 3' UTR



קו וורוד מאונך מציין אתר CG.

### טבלה 11 - אפיון המתילציה בגן DEF

אחוז CNG ממותל	אחוז מתילציה בשני אתרי CNG		אחוז ה-CG הממותל	אחוז מתילציה בשלושה אתרי CG			מספר אתרי CNG	מספר אתרי CG	פנוטיפ
	1992	1983		2473	2404	1828			
0.32	5.16	2.78	0.2	3.03	3.33	3.33	30	41	WT
0	0	0	0.1	0	3.33	0			Off-type

המקטע הנבדק של הגן DEF מורכב משלושה מקטעים בעלי חפיפה קטנה ביניהם שרוצפו בשיטת Bisulfite sequencing בעצי WT ו-Off-type של 'ברהי' בעלי פנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות' ברקמות עלה ותפרחות. התוצאות המוצגות הן ממוצע של שלושה פרטים מכל פנוטיפ, כאשר בכל פרט רוצפו בין שתיים לשש מושבות עבור כל מקטע. באזורים החופפים בין שלושת המקטעים הנתונים מייצגים ממוצע המבוסס על כ-12-26 מושבות.

## 4.2. אפיון דגם הביטוי בעצי תמר ננסיים ובעצי תמר נורמאליים על ידי

### ריצוף בהספק גבוה

נעשו שתי אנליזות לבחינת דגם הביטוי של עצי תמר נורמאליים וננסיים בשיטת הריצוף של Solexa/Illumina על ידי חברת FASTERIS בשווייץ. האנליזות שבוצעו הן Digital Gene Expression (DGE) mRNA-seq ושיטת ה-DGE מבוססת על יצירת תגים קצרים ייחודיים עבור כל גן וריצופם בהספק גבוה. תגים אלו מייצגים בדרך כלל את האזור ה-3' של הגן. השיטה מאפשרת לבחון את רמות הביטוי של הגנים על ידי חישוב כמות התגים המזוהה עבור כל גן (Feng *et al.*, 2010). בשיטת ה-mRNA-seq נעשית פרגמנטציה מכאנית של ה-mRNA (המכיל פולי (A) וריצוף של רצפים קצרים (במקרה שלנו נעשה הריצוף ב-mRNA-seq משני קצוות המקטע - השיטה מכונה Paired end).

### 4.2.1. אפיון דגם הביטוי באמצעות Digital Gene Expression

#### 4.2.1.1. תוצאות ה-Digital Gene Expression

נעשתה השוואה בין דגמי הביטוי של "לבבות" משני עצים: עץ תמר נורמאלי ועץ ננסי. התקבלו רשימות של תגים שנוצרו על ידי חיתוך בחמישה אנזימי רסטריקציה *HinP1I*, *NlaIII*, *TaqI*, *HpaII*, *HpyCH4IV*. סה"כ רוצפו 4,431,077, 1,167,063, 821,114, 410,431 ו-1,404,985 תגים קצרים בכל אנזים בהתאמה. בכל אנזים נקבע מספר הפעמים שכל תג הופיע בעץ הנורמאלי ובעץ הננסי. התוצאות נורמלו לפי robust quantile (Bolstad *et al.*, 2003). נירמול זה מביא להורדת השונות שאינה נובעת ממקור ביולוגי. בנוסף, התגים עברו תיקון Bonferroni ל"ריבוי מבחנים". התגים סוננו לערך טרנספורמציה גבוה משש (הערך מיצג תוצאות בעלות התפלגות נורמאלית) וערך p נמוך מ-0.001 (פרק 3.2.5 בשיטות וחומרים). מספר התגים השונים שנותר לאחר הסינון היה: 1604, 500, 123, 135 ו-511 עבור *HinP1I*, *NlaIII*, *HpaII*, *HpyCH4IV* ו-*TaqI* בהתאמה (טבלה 12). בחנו תגים בעלי ביטוי שונה בין עץ WT ל-DW, ביחס של לפחות פי שלוש בין WT ל-DW. לאחר סינון זה כמות התגים שנשארה הייתה 627, 155, 36, 48 ו-172 עבור *HinP1I*, *NlaIII*, *HpyCH4IV*, *HpaII* ו-*TaqI* בהתאמה (טבלה 12).



## טבלה 12 - תגים השונים בדגם הביטוי בין WT ל-DW

<i>TaqI</i>		<i>HpaII</i>		<i>HpyCH4I</i>		<i>HinPII</i>		<i>NlaIII</i>		אנזים רסטריקציה	
CGA		CGG		CGT		CGC		CATG		אתרי החיתוך	
%	511	%	135	%	123	%	500	%	1604	*מספר הרצפים שעברו סינון	
56.4	97	60.4	29	38.9	14	41.3	64	50.9	319	wt/dw <0.3	רצפים בעלי דפוס ביטוי שונה בשני הפנוטיפים
43.6	75	39.6	19	61.1	22	58.7	91	49.1	308	wt/dw >3	
<b>100</b>	<b>172</b>	<b>100</b>	<b>48</b>	<b>100</b>	<b>36</b>	<b>100</b>	<b>155</b>	<b>100</b>	<b>627</b>	סה"כ	

\* הרצפים עברו סינון לערך טרנספורמציה גבוה ממש וערך p נמוך מ-0.001 לאחר תיקון Bonferroni (ראה פרק 3.2.5 בשיטות וחומרים).

### 4.2.1.2. הארכת רצף באמצעות 5' RACE

מכיוון שהתגים שהתקבלו מאנליזת Solexa היו קצרים מאוד, נבחרו 100 תגים שריצפם הווארך בשיטת 5' RACE על ידי חברת Fasteris בשווייץ. התגים נבחרו כדלקמן:

- 70% בעלי היחס הגבוה ביותר או הנמוך ביותר בין WT ל-DW.
- 15% בעלי יחס קרוב לפי 3 או 0.33 בין WT ל-DW.
- 15% בעלי יחס לפחות פי 3 או 0.33 בין WT ל-DW.

ביניהם נבחרו תגים בהם התקבלה תוצאה חיובית בנייתו ה-Blast. מתוך 100 התגים הנ"ל ההארכה הצליחה ב-58 מהתגים. בחלק מהמקרים הריצוף נעשה על תוצרי ההארכה המקוריים ובמקרים אחרים התקבלו כמה תוצרים ששובטו לתוך פלסמידים ורוצפו או שהתקבל smear. 24 תגים הווארכו כתוצר אחד (או שהתקבלו רצפים דומים במספר תוצרים).

### 4.2.1.3. אנליזת BLAST

#### 4.2.1.3.1. אנליזת BLAST כללית

אנליזת Blast כללית נעשתה ל-24 הרצפים שהווארכו בשיטת ה-5' RACE. האנליזה התבססה על השוואה בין רצף הנוקליאוטידים בבסיס נתונים של nr/nt ו-ESTs. התקבלו תוצאות חיוביות ב-12 רצפים. נבחרו רק תוצאות עם Max Identity (ערך המכמת את מידת ההומומולוגיה בין שני הרצפים) גבוהה מ-83% ו-E value (ערך שמתאר את הסיכוי שההומומולוגיה הינה אקראית) נמוך מ- $9E-47$  (כלומר נמוך מ- $9 \cdot 10^{-47}$ ). התוצאות מסוכמות בטבלה 13.

#### 4.2.1.3.2. אנליזת BLAST לגנום התמר

ה-Blast נעשה מול טיוטת גנום התמר (שפורסמה במאי 2009). נמצאו 19 רצפים בעלי הומומולוגיה לגנום. התוצאות מסוכמות בטבלה 14. שבעה מתוכם היו בעלי הומומולוגיה גבוהה. שנים מהרצפים האלה (Sol 5 ו-Sol 12) הראו הומומולוגיה גבוהה לשני Scaffolds (קטעים) שונים בטיוטת הגנום.

12 מהרצפים היו בעלי הומוולוגיה נמוכה יותר לטיוטת הגנום. בחמישה גנים (Sol 7, Sol 517, CGG 135, CGC 420 ו-CGG 134) לא נמצאה הומוולוגיה של רצף ה-RACE 5' לטיוטת גנום התמר.

**טבלה 13 - תוצאות BLAST כללי של הרצפים שהוארכו ב-5' RACE**

EST				nr/nt				שם הרצף	מספר
E value*	Max Identity	Accession	אורגניזם	E value*	Max Identity	Accession	אורגניזם		
3E-173	98%	ES414513.1	<i>E. guineensis</i>	1E-171	98%	AF295636.1	<i>E. guineensis</i>	Sol 3	1
9E-128	94%	EL686308.1	<i>E. guineensis</i>	1E-127	94%	EU284857.1	<i>E. guineensis</i>	Sol 5	2
2E-57	97%	EY396333.1	<i>E. guineensis</i>	4E-49	93%	EU284940.1	<i>E. guineensis</i>	Sol 6	3
9E-58	85%	EL681295.1	<i>E. guineensis</i>				לא נמצאה הומוולוגיה	Sol 7	4
1E-106	92%	EL689458.1	<i>E. guineensis</i>	2E-78	86%	CT841600.1	<i>O. rufipogon</i>	Sol 11	5
2E-119	95%	EL683809.1	<i>E. guineensis</i>	6E-65	86%	EU284871.1	<i>E. guineensis</i>	Sol 12	6
3E-122	94%	EL693820.1	<i>E. guineensis</i>	5E-111	96%	EU285015.1	<i>E. guineensis</i>	Sol 17	7
2E-155	93%	EL681222.1	<i>E. guineensis</i>	6E-82	85%	AY389652.1	<i>H.s orientalis</i>	Sol 19	8
9E-47	83%	GH197794.1	<i>P. fascicularis</i>				לא נמצאה הומוולוגיה	CGC 466	9
5E-74	90%	EL693758.1	<i>E. guineensis</i>				לא נמצאה הומוולוגיה	CGC 495	10
2E-57	86%	ES323860.1	<i>E. guineensis</i>	4E-55	86%	FJ940771.1	<i>E. guineensis</i>	CGC 499	11
לא נמצאה הומוולוגיה				5E-94	100%	X76545.1	<i>P.dactylifera</i>	CGG 135	12

ל-12 הרצפים לא נמצאה הומוולוגיה בשני המאגרים EST ו-nr/nt: Sol 1, Sol 8, Sol 14, Sol 291, Sol 503, Sol 510, Sol 517, CGC 420, CGC 438, CGG 2, CGG 134 ו-CGT 123. \* הערכים כתובים בצורת XE-Y שהיא מקבילה ל- $10^{-Y}$ .

טבלה 14 - תוצאות BLAST לגנום התמר של הרצפים שבודדו ב-5' RACE

תוצאות BLAST עבור רצפים שהארכו ב-5' RACE				שם הגן	מספר
E value	Max Identity	רצף גנומי בטיטת התמר	אורך רצף ה-5' RACE		
E-428	86/86 (100%)	ACYX01050427.1	103	Sol 1	1
E-125	232/234 (99%)	ACYX01011492.1	351	Sol 3	2
E-561	111/111 (100%)	ACYX01050427.1	127	Sol 8	3
3E-52	112/116 (96%)	ACYX01081795.1	205	CGC 499	4
E-146	279/288 (96%)	ACYX01103836.1	288	Sol 5	5
3E-34	167/200 (83%)	ACYX01037884.1			
4E-29	65/65 (100%)	ACYX01125861.1	136	Sol 6	6
1E-67	160/175 (91%)	ACYX01127074.1	253	Sol 11	7
1E-64 ,1E-67	125/125 (100%) ,130/130 (100%)	ACYX01091764.1	275	Sol 12	8
8E-23 ,2E-38	85/95 (89%) ,114/125 (91%)	ACYX01022983.1			
E-156	282/283 (99%)	ACYX01107628.1	307	Sol 14	9
8E-17 ,2E-51	45/45 (100%) ,103/103 (100%)	ACYX01065725.1	300	Sol 17	10
E-134	245/246 (99%)	ACYX01071343.1	440	sol 19	11
6E-81	155/156 (99%)	ACYX01095170.1	169	Sol 503	12
8E-77	145/145 (100%)	ACYX01109303.1	157	Sol 510	13
E-107	203/205 (99%)	ACYX01023238.1	510	Sol 291	14
3E-79	152/155 (98%)	ACYX01037310.1	155	CGC 438	15
5E-11 ,2E-22 ,E-39	35/35 (100%) ,54/54 (100%) ,83/83 (100%)	ACYX01015434.1	217	CGC 495	16
7E-72	137/137 (100%)	ACYX01063695.1	204	CGC 466	17
2E-21	52/52 (100%)	ACYX01050427.1	101	CGG 2	18
2E-90	174/176 (98%)	ACYX01111370.1	189	CGT 123	19

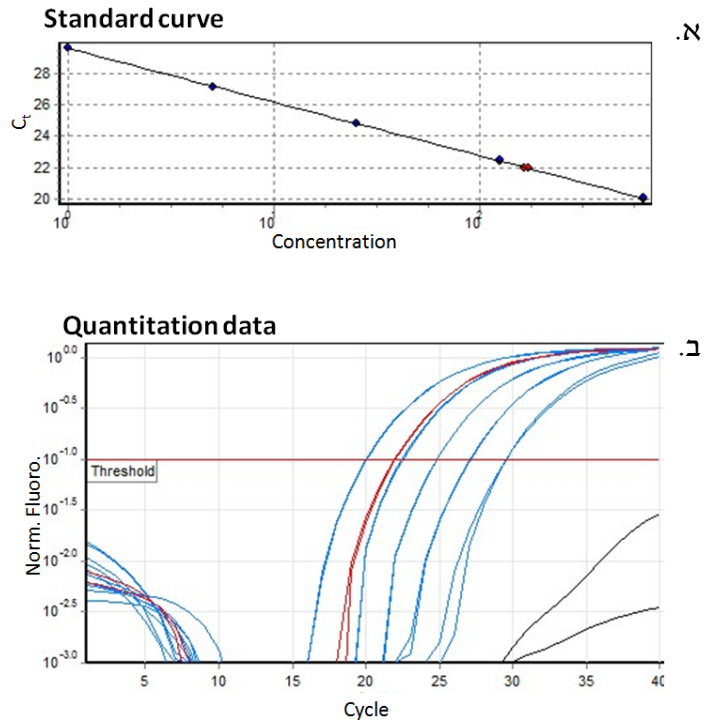
עבור הרצפים הבאים: Sol 7, Sol 517, CGG 135, CGG 420, CGG 134-1 לא נמצאה הומולוגיה בין הרצף מ-5' RACE לטיטת הגנום.

#### 4.2.1.4. אימות דגם הביטוי של ה-DGE באמצעות Q-RT-PCR

בכדי לאמת את דגם הביטוי שהתקבל באנליזת ה-DGE, נעשתה בדיקה באמצעות Q-RT-PCR. התבנית לבדיקת Q-RT-PCR הוכנה מ-RNA שהופק מרקמת "לב" התמר (ראה שיטות וחומרים, איור 1) ושימש להכנת cDNA. לכל אחד מ-24 הרצפים שהוארכו ב-RACE 5' תוכננו 2-3 זוגות תחלים ייחודיים בתוכנות Primer Express ו-Primer 3 (טבלה 7) ונבחר זוג התחלים שעמד בכל הקריטריונים הבאים: נתנו תוצר אחד בריאקציית PCR כמותית, התוצר היה בגודל הנכון והיה בעל הומולוגיה לרצף ממנו הוא סונתז ונתן עקומת כיוול טובה ב-Q-RT-PCR עם תוצר אחד בלבד. רק ל-12 רצפים מתוך 24 הצלחנו לסנתז תחלים שעמדו בכל הקריטריונים הנ"ל. (עבודה זאת נעשתה בשיתוף פעולה עם ד"ר Hewan Demissie Degu – פוסט דוקטורנטית במעבדה). אני ניתחתי 10 מ-24 הרצפים ומתוכם 7 נתנו תוצאות ב-Q-RT-PCR. מקטע של הגן הריבוזומאלי 18S שבודד מתמר נורמאלי על סמך הרצף של הגן האורתולוגי מדקל השמן שימש כגן ייחוס (reference). בדיקת רמת ביטוי של הרצפים שנבחרו, נעשתה בהשוואה לדגם הביטוי של הגן 18S. לכל הגנים שנבדקו נבנו עקומות כיוול מתוך ערכי ה- $C_t$  (threshold cycle) כנגד לוגריתם של ריכוז ה-cDNA (איור 9), וחושב ערך Relative Expression (ראה חומרים ושיטות).

דגם הביטוי שנמצא באנליזה ה-Q-RT-PCR לא היה דומה בשום מקרה לדגם שהתקבל באנליזה ה-DGE. לשלושה רצפים (מתוך 10 שאני בחנתי) לא הצלחתי לקבל עקומת כיוול ב-Q-RT-PCR. בשישה רצפים נתקבל ב-Q-RT-PCR יחס קרוב ל-1:1 או יחסים של עד פי שלוש בין העץ הנורמאלי לננס. לעומת זאת באנליזת ה-DGE ארבעה מתוך הרצפים הנ"ל נתנו יחס מאוד גבוה בין עצי WT לעצי DW ושני רצפים נתנו יחס מאוד נמוך. רק ברצף אחד (CGT123) התקבל ערך Relative expression של 0.3, (כלומר ביטוי של הגן הזה בעץ נורמאלי הוא נמוך פי שלוש מאשר הביטוי של הגן בעץ הננסי) וגם זאת שלא בהתאמה לתוצאות של ה-DGE בהן היחס היה 0.12 (טבלה 15). בארבעה מהמקרים בחנתי שני זוגות תחלים עבור כל רצף. במקרים אלה שני הזוגות נתנו יחס דומה בין עץ WT ל-DW. הרצפים שנבחנו ע"י ד"ר Hewan Demissie Degu נתנו גם כן יחס הקרוב ל-1:1 בין העץ הנורמאלי לננס באנליזת Q-RT-PCR למרות שבאנליזת DGE נתקבלו יחסים של מעל לשלוש או מתחת ל-0.33. רק באחד מהרצפים שהיא בחנה (Sol 19) התקבל יחס של פי שלוש בין העץ הנורמאלי לננס, זה לא תואם את תוצאות ה-DGE בהן היחס בין העץ הנורמאלי לננס היה 50.

איור 9 - כיול ואנליזה של Q-RT-PCR לגן 8 Sol



- א. עקומת כיול ליניארית (Standard curve). ציר ה-X: לוגריתם של ריכוז cDNA שחושב ככמות (במיקרוליטרים) מריאקציית ה-cDNA. ציר ה-Y: ערכי C<sub>t</sub>.
- ב. אנליזה כמותית - ציר ה-Y: רמת פלואורסנציה (המבטאת את כמות התוצר המסונתזת). ציר ה-X: מספר המחזורים של ריאקציית ה-PCR. הקו האדום המאונך מייצג את קו ה-Threshold שחושב על ידי התוכנה ובעזרתו מחושבים ערכי ה-C<sub>t</sub>. חמשת העקומות הכחולות בגרף מייצגות ריכוזים שונים של cDNA התחלתי של עצי WT (המייצגים את עקומת הכיול של הגן) והעקומה האדומה מייצגת את ה-cDNA של פרט off-type. העקומות השחורות מייצגות דוגמאות ללא cDNA.

טבלה 15 - רמות ביטוי של גנים על פי DGE ועל פי Q-RT-PCR בעץ נורמאלי וננס

יחס ביטוי הגנים באנליזה Q-RT-PCR				יחס ה-WT/DW לפי אנליזה DGE	שם הרצף	מספר
ממוצע הביטוי בתחלים שונים לאותו גן	זוג תחלים 3	זוג תחלים 2	זוג תחלים 1			
0.82	NA	0.939	0.703	720.38	Sol 1	1
1.8	1.8	NA	NA	127.84	Sol 5	2
0.72	0.69	NA	0.75	99.61	Sol 8	3
1.39	NA	1.39	NA	71.32	Sol 12	4
0.78	NA	0.6	0.95	0.33	CGC 438	5
0.33	NA	0.357	0.31	0.12	CGT 123	6
0.49	NA	NA	0.49	0.08	Sol 510	7
	NA	NA	NA	78.73	Sol 11	8
	NA	NA	NA	3.1	Sol 291	9
	NA	NA	NA	0.03	CGG 134	10

NA - לא התקבלו תוצרים בריאקציית ה-PCR. יחס הביטוי לכל זוג תחלים וגן הוא ממוצע של 2 חזרות. לכל רצף סונתזו שלושה זוגות תחלים. לא מופיעות בטבלה התוצאות עבור 14 רצפים נוספים אשר נבחנו ע"י ד"ר Hewan Demissie Degu.

## 4.2.2. אפיון דגם הביטוי באמצעות mRNA-seq

### 4.2.2.1. תוצאות ה-mRNA-seq

אנליזת ה-mRNA-seq נעשתה על תערובת RNA מ-15 "לבבות" של עצים ננסים (DW) ותערובת משני "לבבות" של עצים נורמאליים (WT) (טבלה 5) בהבדל מאנליזת ה-DGE שנעשתה על עץ WT אחד ועץ ננס אחד. האנליזה נעשתה בשיטת mRNA-seq על ידי חברת Fasteris. התקבלו שני דגמי ביטוי, אחד עבור תערובת WT והשני עבור תערובת DW. בשיטת ה-mRNA-seq מרוצפים 36 בסיסים מכל כיוון של מקטע ה-DNA (שלושת הבסיסים הראשונים היו ה"חותמת" שמזהה כל אחת מהתערובות) כל דגם ביטוי מורכב מקריאות משני צידי "מושבת ה-DNA" שרוצפה. תערובת העצים הנורמאליים הכילה בסה"כ 13,647,010 קריאות ותערובת העצים הננסים הכילה בסה"כ 14,443,018 קריאות. בסה"כ נכללו באנליזה 450,351,330 בסיסים בעצים הנורמאליים ו-476,619,594 בסיסים בעצים הננסים (טבלה 16).

### טבלה 16 - מספר הרצפים שהתקבלו באנליזת mRNA-seq

שם הדוגמא	כיוון קריאה	*מספר קריאות	מספר בסיסים
תערובת WT	כיוון א	6,823,505	225,175,665
	כיוון ב	6,823,505	225,175,665
	סה"כ	13,647,010	450,351,330
תערובת DW	כיוון א	7,221,509	238,309,797
	כיוון ב	7,221,509	238,309,797
	סה"כ	14,443,018	476,619,594

כל מקטע באנליזה רוצף משני הכיוונים. \* כל 33 בסיסים מהווה קריאה אחת.

### 4.2.2.2. סינון התוצאות

לרצפים שהתקבלו מאנליזת mRNA-seq נעשה assembly על ידי ד"ר רון אופיר באמצעות תוכנת SOAP. ההשוואה נעשתה ל-19,414 רצפים מטיוטת הגנום של התמר, שלפי אנליזה ביואינפורמטית צפויים להתבטא כרצף המקודד לחלבון כפי שהוגדר באתר הבא: <http://qatar-weill.cornell.edu/research/datepalmGenome/download.html>. מתוך הרצפים שהתקבלו ב-mRNA-seq (ראה טבלה 16) מופו במהלך זה 17,032 רצפים. הרצפים עברו מספר סינונים:

1. תיקון FDR (False discovery rate), המייצג יחס של ערכים מצופים כלפי השגיאות והמתקן את השגיאות האקראיות כתוצאה מריבוי המבחנים).

2.  $\text{Log}_{\text{Con}}$  (ממוצע לוג של ריכוז הרצפים הספציפיים משתי הקבוצות ביחס לכלל הרצפים האחרים; הערך מאפשר לסנן רק רצפים בעלי מספר קריאות גבוה).

ה-FDR סונן לערכים קטנים מ-0.01, 0.001 ו-0.0001 בנוסף נעשה סינון ל- $\text{Log}_{\text{Conc}}$  נמוך מ-17 (ערך המשאיר רצפים בעלי מספר גדול של קריאות ומסנן רצפים שזוהו מעט פעמים בדוגמא). בנוסף נעשה סינון לביטוי שונה בין תערובת עצי WT ל-DW, ביחס של לפחות פי 3 בין WT ל-DW. תוצאות הסינון מסוכמות בטבלה 17.

טבלה 17 - רצפים שנבדלו בדגם הביטוי בין WT ל-DW

<i>FDR</i> <0.0001		<i>FDR</i> <0.001		<i>FDR</i> <0.01		סינון ברמות מובהקות שונות של <i>FDR</i>	
ללא סינון של $\text{Log}_{\text{Conc}}$							
מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)	
%	4577	%	5027	%	5540	wt/dw <0.33	רצפים בעלי דפוס ביטוי שונה בעץ נורמאלי וננס סה"כ
74.5	1458	75.5	1570	76	1685	wt/dw >3	
25.5	498	24.5	509	24	532	סה"כ	
100	1956	100	2079	100	2217		
<i>FDR</i> ≤0.0001		<i>FDR</i> ≤0.001		<i>FDR</i> ≤0.01		סינון ברמות מובהקות שונות של <i>FDR</i>	
עם סינון של $\text{Log}_{\text{Conc}}$							
מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)		מספר הרצפים שעברו סינון (filter)	
%	4300	%	4693	%	5121	wt/dw <0.33	רצפים בעלי דפוס ביטוי שונה בעץ נורמאלי וננס סה"כ
75.2	1295	76.1	1363	76.6	1418	wt/dw >3	
24.8	427	23.9	429	23.4	433	סה"כ	
100	1722	100	1792	100	1851		

תוצאות אנליזת SOAP סוננו ב-3 רמות מובהקות שונות של *FDR* עם וללא סינון של  $\text{Log}_{\text{Conc}}$ . בנוסף הערכים סוננו לבעלי דגם ביטוי של פי 3 לפחות בין עצים נורמאליים וחריגים.

**4.2.2.3. אימות דגם הביטוי של ה-mRNA-seq באמצעות Q-RT-PCR**

בכדי לאמת את דגם הביטוי שהתקבל באנליזת mRNA-seq נעשתה בדיקה ב-Q-RT-PCR בשני שלבים. בשלב הראשון הבדיקה נעשתה על תערובת עצי DW שהכילה את כל אותם העצים שנבחנו ב-mRNA-seq מול עץ WT אחד. בשלב השני נבחרו מספר גנים מהשלב הראשון ובהם אנליזת ה-Q-RT-PCR נעשתה על פרטים בודדים: ששה עצי DW (שהיו חלק מהתערובת בשלב הראשון) ושני עצי WT שהיו בתערובת באנליזת mRNA-seq.

בשלב הראשון התבנית לבדיקת Q-RT-PCR הוכנה מ-RNA של תערובת DW ו-RNA מעץ WT בודד (עקב מחסור ברקמה של ה-WT השני) שהופק מרקמת "לב" התמר (ראה חומרים ושיטות). נבחנה רמת הביטוי של 24 רצפים המייצגים מגוון רחב של יחסים שונים של רמות ביטוי בין תערובת העצים נורמאליים לחריגים. טווח יחסי הביטוי בין ה-WT וה-DW של הגנים השונים שנבחנו ב-Q-RT-PCR היה (לפי אנליזת ה-mRNA-seq) כ-85,000. חלק מרצפים שנבחרו צפויים להיות קשורים לצימוח ווגטטיבי או כאלה שמוטנטים שלהם בצמחים אחרים קשורים לננסות. אחד מהרצפים היה בעל יחס קרוב לאחד בין עצים נורמאליים לחריגים ושימש כביקורות. 20 מהרצפים נבחרו לאחר סינון  $\text{Log}_{\text{Conc}} \geq -17$  ו- $\text{FDR} < 0.0001$  (בעלי מספר קריאות גדול ולכן בעלי הבדלים סיגניפיקנטיים), ארבעה מהרצפים נבחרו לאחר סינון  $\text{FDR} < 0.0001$  ולא עברו סינון  $\text{Log}_{\text{Conc}} \geq -17$  (רצפים אלה נבחרו עקב יחסי הביטוי הקיצוניים שלהם בין WT ל-DW). לכל רצף תוכנו 2-3 זוגות תחלים ייחודיים בתוכנות Primer Express ו-Primer 3 (טבלה 8).

מתוך 11 רצפים בעלי יחס גדול מפי שלוש בין WT ל-DW באנליזת mRNA-seq, תשעה היו בעלי יחס גדול מפי שלוש בין WT ל-DW גם באנליזת Q-RT-PCR. ככל שהיחס בין עצים נורמאליים לננסים היה גבוה יותר באנליזת mRNA-seq כך הוא היה גבוה יותר גם באנליזת Q-RT-PCR. בשני רצפים נמצא יחס

של קרוב לאחד ובהם לא אומתו תוצאות ה-mRNA-seq. 12 מהרצפים שנבחרו היו בעלי יחס נמוך מפי שלוש בין WT ל-DW באנליזת mRNA-seq. רצפים אלה הראו גם יחס דומה באנליזת Q-RT-PCR. ככל שהיחס ברמת הביטוי בין תערובת של עצים נורמאליים לננסיים היה גדול יותר באנליזת mRNA-seq הוא היה גדול יותר גם באנליזת Q-RT-PCR. הגן ששימש לביקורת והיה בעל יחס דומה לאחד (0.9) מאנליזת mRNA-seq הראה יחס של 0.4 בין עצים נורמאליים לננסיים. רוב התוצאות אומנם לא חזרו במדויק על היחס שהתקבל באנליזת mRNA-seq אבל הראו הבדלים משמעותיים ברמת הביטוי בין עצי WT לעצי DW התואמים בכיוונם את תוצאות אנליזת ה-mRNA-seq. התוצאות מסוכמות בטבלה 18.

**טבלה 18 - רמות ביטוי של גנים על פי mRNA-seq ו-Q-RT-PCR בתערובות של עצים נורמאליים וננסיים**

ממוצע WT/DW לפי אנליזת Q-RT- (±) PCR (ט.ס)	יחס ביטוי בחזרות של Q-RT-PCR				WT/DW לפי אנליזת mRNA- seq	תפקיד צפוי לגן לפי הומולוגיה לרצפים אחרים (Annotation)	שם הרצף	מספר
	4	3	2	1				
11.4±6.26	6.8	19	14	5.8	236/0	1 aminocyclopropane 1 carboxylate oxidase	PDK_20s1300981g0 01	1
69.4±11.09			61.6	77.3	349.5	cationic peroxidase 1	PDK_20s1955911g0 01	2
11.7±0.1			11.74	11.6	148.6	helix loop helix dna- binding domain containing protein	PDK_20s1685261g0 01	3
11.9±5.8			7.8	16	133	WRKY transcription factor	PDK_20s1368561g0 02	4
1.6±0.53			1.22	2	39	myo-inositol oxygenase	PDK_20s1530771g0 01	5
11.7±0.42			12	11.4	27.25	1 xyloglucan endotransglycosylase xet2	PDK_20s1959831g0 01	6
6.7±2.97			4.6	8.8	27	gibberellin 20	PDK_20s1306264g0 01	7
4.8±0.49			5.19	4.5	17.8	Fe oxygenase family expressed	PDK_20s1301376g0 02	8
16.4±3.36			14	18.8	14	expansin-like protein a	PDK_20s1398571g0 01	9
3.7±0.05			3.7	3.7	13.8	legumin-like protein	PDK_20s1325401g0 01	10
1.47±0.13			1.4	1.6	10.5	3 -n-debenzoyltaxol n-benzoyltransferase -like	PDK_20s1377181g0 01	11
0.38±0.03			0.355	0.4	0.9	ap2 erf domain- containing transcription factor	PDK_20s1619651g0 01	12
0.11			0.11	0.11	0.24	similar to predicted protein	PDK_20s1359671g0 01	13

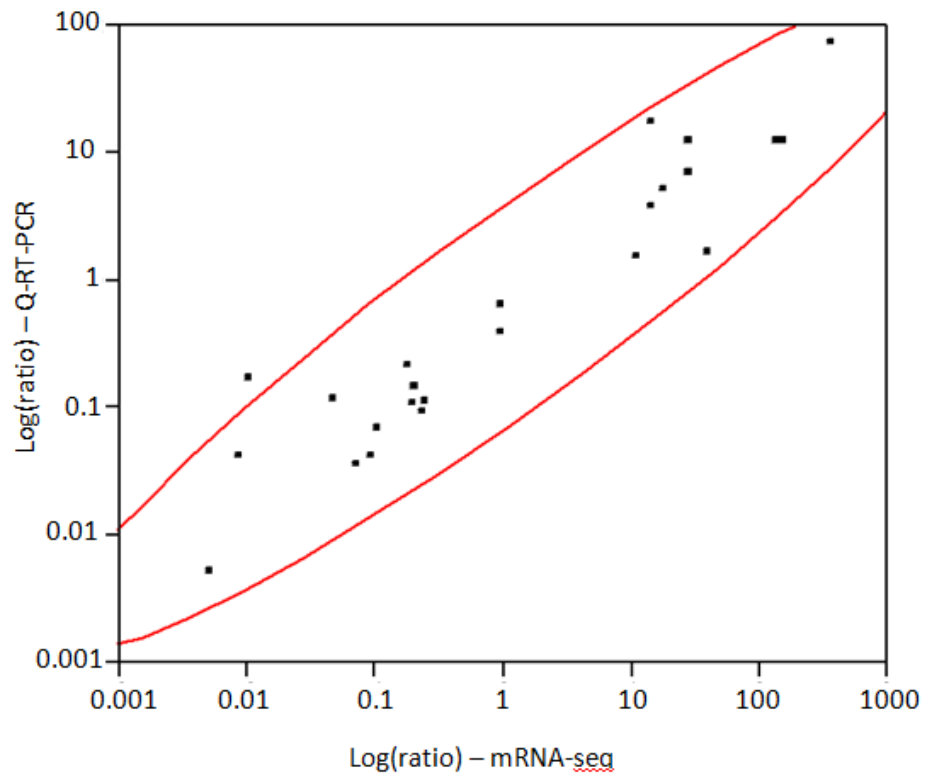


ממוצע WT/DW לפי אנליזת Q-RT- (±) PCR (ת.ס)	יחס ביטוי בחזרות של Q-RT-PCR				WT/DW לפי אנליזת mRNA- seq	תפקיד צפוי לגן לפי הומולוגיה לרצפים אחרים (Annotation)	שם הרצף	מספר
	4	3	2	1				
0.09±0.01			0.10	0.08	0.23	auxin response factor 8	PDK_20s1359912g002	14
0.14±0.05			0.10	0.18	0.20	auxin response transcription factor	PDK_20s1699311g001	15
0.1±0.06			0.15	0.06	0.19	brassinosteroid-insensitive 1	ACYX01096809.1	16
0.2±0.01			0.21	0.20	0.18	brassinosteroid-insensitive 1	PDK_20s1439951g001	17
0.07±0.005			0.07	0.06	0.10	auxin response factor	PDK_20s1674351g001	18
0.04			0.04	0.04	0.09	auxin response factor 7a	PDK_20s1410001g002	19
0.035±0.01			0.03	0.04	0.07	glycosyl hydrolase family 9	PDK_20s1743471g001	20
0.11±0.08		0.08	0.2	0.07	0.05	cytokinin oxidase	PDK_20s1479321g001	21
0.04±0.03		0.02	0.07	0.03	0.0085	dwarfl diminuto	ACYX01151934.1	22
0.17±0.05			0.13	0.20	0/54	similar to predicted protein	PDK_20s1454381g003	23
0.005±0.001			0.006	0.004	0/61	brassinosteroid insensitive 1-associatedreceptor kinase 1	PDK_20s1392771g001	24

הגנים בטבלה מסודרים לפי ערכים יורדים של יחס WT/DW מאנליזת mRNA-seq. באותם מיקרים בהם רק באחת מהתערובות התקבל ריצוף של הקונטיג, מוצג היחס של הרצפים כשבר בו המונה או המכנה הוא אפס. "ס.ת" הינה סטיית תקן.

הקורלציה בין תוצאות אנליזת ה-mRNA-seq וה-Q-RT-PCR נבחנה בתוכנת JMP. מכיוון שטווח היחסים היה רחב מאוד, נעשתה הקורלציה לערכים הלוגריתמיים של יחסי הביטוי השונים לכל גן וגן בעצי WT כנגד עצי DW. גרף המייצג את ההשוואה של שתי האנליזות מוצג באיור 10 (לצורך בניית הגרף הושמטה נקודה אחת (גן PDK\_20s1300981g001), בה לא התאפשרה הטרנספורמציה הלוגריתמית מכיוון שהיה בה חילוק ב-0). האנליזה הצביעה על קורלציה טובה בין שתי השיטות, עם ערך קורלציה (R) של 0.944 וערך מובהקות (p) של  $4.85E-12$ . חשוב לציין שבדרך כלל ההבדלים ביחס הביטוי באנליזת ה-mRNA-seq היו גבוהים יותר מאלה שהתקבלו באנליזת ה-Q-RT-PCR עם מובהקות סטטיסטית של 0.0022.

## איור 10 - קורלציה בין תוצאות דגם הביטוי מאנליזות mRNA-seq וה-Q-RT-PCR



הקורלציה נעשתה על 24 גנים. האליפסה האדומה מייצגת מובהקות של 95%.

בשלב הבא בחנתי את רמת הביטוי של ארבעה גנים בשני עצים נורמאליים ובשישה עצים חריגים בנפרד (לא בתערובות). נבחרו גנים שנמצאו להיות בעלי הבדלים קיצוניים ברמת הביטוי בין תערובת WT לתערובת DW. שנים מתוך הגנים היו בעלי יחס מאוד נמוך בין WT ל-DW ושניים אחרים בעלי יחס גבוה מאוד. ארבעת הגנים נבדקו בשמונת העצים באותה הרצה. לאחר אנליזת Q-RT-PCR כל פרט של כל גן נורמל לגן הביקורת (18S) וחושב ממוצע הריכוזים של עצי WT ו-DW עם שגיאות התקן שלהן. כמו כן חושב גם היחס בין ממוצע עצי ה-WT לממוצע עצי ה-DW. התוצאות מוצגות בטבלה 19. בכל הגנים נמצא הבדל בולט בביטוי בין ממוצע הפרטים החריגים וממוצע העצים הנורמאליים. רמת הביטוי של שלושת הגנים PDK\_20s1955911g001, ACYX01151934.1 ו-PDK\_20s1392771g001 דומה בשני עצי ה-WT, אבל בגן PDK\_20s1685261g001 רמת הביטוי שונה מאוד (הבדל של כמעט פי 13 בין שני הפרטים). בבחינה של כל גן בנפרד נמצא שקיימת שונות בתוך קבוצת ה-Off-type בין הפרטים (טבלה 19). בשלושת הגנים PDK\_20s1955911g001, PDK\_20s1685261g001 ו-ACYX01151934.1 יש קורלציה טובה יותר בין אנליזת mRNA-seq לאנליזת Q-RT-PCR של פרטים בודדים. ערך הקורלציה בין שלושת האנליזות נבחן בתוכנת JMP. מקדם הקורלציה בין אנליזת התערובות לממוצעי האנליזה של הפרטים הבודדים ב-Q-RT-PCR הוא 0.989. מקדם הקורלציה בין אנליזת mRNA-seq לממוצע האנליזת Q-RT-PCR של פרטים בודדים הוא 0.992 ומקדם הקורלציה בין אנליזת mRNA-seq לאנליזת Q-RT-PCR של תערובות ה-RNA הוא 0.964

(טבלה 20). ערכי המובהקות באנליזה זאת אינם גבוהים ונעים מ-0.008 ועד 0.036 (טבלה 20), כלומר מובהקים רק ברמה של  $p < 0.05$ , וזאת כנראה עקב המגבלה של מספר הגנים המועט שנבחנו.

**טבלה 19 - ביטוי של ארבעה גנים בעצי WT ו-DW בודדים**

PDK_20s139277 1g001	ACYX011519 34.1	PDK_20s168526 1g001	PDK_20s195591 1g001	גן	
0.0006	0.03	3.35	0.26	WT	יחס ביטוי בין הגן ל-18S
0.001	0.01	0.26	0.26	WT2	
<b>0.001±0.0002</b>	<b>0.023±0.007</b>	<b>1.81±1.09</b>	<b>0.26±0.001</b>	ממוצע WT (± ש.ת.)	
0.23	0.232	0.06	0.001	DW2	יחס ביטוי בין הגן ל-18S
0.31	1.239	0.002	0.003	DW3	
0.06	0.67	0.016	0.002	DW9	
0.33	0.48	0.065	0.001	DW13	
0.003	2.58	0.006	0.001	DW14	
0.13	1.9	0.05	0.001	DW16	
<b>0.177±0.02</b>	<b>1.18±0.15</b>	<b>0.03±0.005</b>	<b>0.0015±0.0001</b>	ממוצע DW (± ש.ת.)	
<b>0.005</b>	<b>0.02</b>	<b>53.88</b>	<b>175</b>	ממוצע WT/DW	
<b>0.005</b>	<b>0.04</b>	<b>11.7</b>	<b>69.4</b>	WT/DW לפי אנליזת Q-RT-PCR של תערובות עצי DW ו-WT	
<b>0/61*</b>	<b>0.0085</b>	<b>148.6</b>	<b>349.5</b>	WT/DW לפי אנליזת mRNA-seq של תערובות עצי DW ו-WT	

המספרים מבטאים את יחס הביטוי של כל גן יחסית לביטוי של 18S. התוצאות מבוססות על חזרה אחת של אנליזת Q-RT-PCR. ש.ת. הינה שגיאת תקן. \* באנליזה של mRNA-seq גן זה לא התבטא ב-WT ונמצאו 61 קריאות שלו בתערובת עצים ננסים.

**טבלה 20 - ערכי הקורלציה בין שלושת האנליזות בארבעת הגנים הנבחרים**

ערך מובהקות	מקדם הקורלציה (R)	משתנה ב'	משתנה א'
0.011	0.989	DW/WT לפי Q-RT-PCR של פרטים בודדים	DW/WT לפי Q-RT-PCR של תערובות
0.008	0.992	DW/WT לפי Q-RT-PCR של פרטים בודדים	DW/WT לפי mRNA-seq
0.036	0.964	DW/WT לפי Q-RT-PCR של תערובות	DW/WT לפי mRNA-seq

ערכי קורלציה בין שלושת האנליזות נבחנו בתוכנת JMP. כל פעם נבחנו מקדם קורלציה בין 2 אנליזות הרשומות תחת משתנה א' ומשתנה ב'.

### 4.2.3. השוואה בין דגם הביטוי לפי DGE ו-mRNA-seq

נעשתה השוואה ראשונית בין שני דגמי הביטוי שהתקבלו מ-DGE ו-mRNA-seq. לצורך כך נעשה Blast לכל 12 תוצרי RACE 5' שנבחנו באנליזת Q-RT-PCR מול רצפי הגנים המקודדים בטיטת הגנום של התמר. בתשעה מתוך 12 הרצפים נמצא רצף mRNA צפוי. רמות הביטוי של רצפי ה-mRNA האלה נבחנו בנתונים שהתקבלו מאנליזת mRNA-seq. חושבה רמת הביטוי של רצפים אלה בתערובות העצים הנורמאליים והחריגים. נעשתה השוואה של רמת הביטוי שהתקבלה באנליזת ה-DGE, אנליזת ה-mRNA-seq ואנליזת ה-Q-RT-PCR (לרצפים שהורארכו ב-RACE 5'). בשבעה מהמקרים היחסים שהתקבלו באנליזת mRNA-seq היו קרובים לאחד. באותם רצפים היחס שהתקבל באנליזת Q-RT-PCR היה גם כן קרוב לאחד ואילו לפי אנליזת DGE, לרצפים אלה יש רמות ביטוי מאוד שונות בין עצי WT ל-DW. ברצף Sol 19 התקבל יחס של פי ארבע באנליזת ה-mRNA-seq, ויחס של פי שלוש באנליזת ה-Q-RT-PCR בין העצים הנורמאליים לחריגים. לעומת זאת באנליזת ה-DGE של Sol 19 התקבל יחס של פי 50 בין עץ WT לעץ ננס (טבלה 21). רצף ה-RACE 5' של Sol 510 לא נתן הומולוגיה לרצפי mRNA. לכן נעשה לו Blast מול רצפי ה-mRNA הצפויים בגנום של התמר ונמצא רצף בעל הומולוגיה גבוהה (Scaffold ACYX01109303.1). לרצף זה נעשה Blast מול רצפי mRNA ונמצא רצף הומולוגי PDK\_20s1695221g001. בין רצף זה לרצף Sol 510 יש מרחק של 145 בסיסים על Scaffold ACYX01109303.1 מטיטת גנום התמר. לכן התייחסתי ל-PDK\_20s1695221g001 כגן זהה ל-Sol 510 בטיטת גנום התמר. רצף Sol 510 נתן יחס ביטוי של 0.1 בין עץ נורמאלי לננס באנליזת ה-mRNA-seq, יחס של 0.49 באנליזת ה-Q-RT-PCR ואילו באנליזת ה-DGE התקבל יחס של 0.08 (טבלה 21).

#### טבלה 21 – השוואה בין האנליזות DGE, mRNA-seq ו-Q-RT-PCR

יחס ה- WT/DW			שם הרצף	מספר
Q-RT-PCR	mRNA seq	DGE		
0.82	0.93	720.38	Sol 1	1
1.13	0.87	150.32	Sol 3	2
1.8	0.49	127.84	Sol 5	3
0.72	0.93	99.61	Sol 8	4
1.39	0.87	71.32	12 Sol	5
0.97	1.16	53.15	17 Sol	6
3	4	50.13	19 Sol	7
0.94	1.25	0.27	466 CGC	8
0.49	0.1	0.08	Sol 510	9*

\* התייחסתי ל-Sol 510 באנליזת DGE ול-PDK\_20s1695221g001 בטיטת גנום התמר המצוי במרחק של רק 145 מרצף ה-TAG כלאותו רצף. הרצפים CGC 438, CGT 123 ו-Sol 14 לא נמצאו בקובץ mRNA ובטיטת גנום התמר.

ערכי הקורלציה והמובהקות בין שלושת האנליזות הושבו בתוכנת JMP. התוצאות מסוכמות בטבלה 22. בין תוצאות אנליזת mRNA-seq לאלה של ה-Q-RT-PCR נמצאה קורלציה של 0.826 עם ערך מובהקות של

0.006. לעומת זאת הקורלציות בין אנליזת DGE ל-mRNA-seq ול-Q-RT-PCR נתנו ערכים גמוכים מאוד אשר אינם מובהקים סטטיסטית.

**טבלה 22 - ערכי הקורלציה בין שלושת האנליזות: DGE, mRNA-seq ו-Q-RT-PCR**

משתנה א'	משתנה ב'	ערך הקורלציה (R)	המובהקות (p)
DGE לפי DW/WT	mRNA-seq לפי DW/WT	-0.110	0.778
mRNA-seq לפי DW/WT	Q-RT-PCR לפי DW/WT	0.826	0.006
DGE לפי DW/WT	Q-RT-PCR לפי DW/WT	-0.169	0.664

ערכי קורלציה בין שלושת האנליזות נקבעו בתוכנת JMP.

## 5. דיון

מטרת המחקר הוא אפיון תהליכי יצירת השונות הסומקלונלית על ידי זיהוי הבדלים בין עצי תמר חריגים מתרבות רקמה לעצים נורמאליים. בעבודה זו נקטנו בשתי שיטות. האחת היא אפיון דגם המתילציה במקטעי DNA מסוימים ובגן DEF, בהם זוהו בעבר הבדלים בין פרטים נורמאליים לחריגים. השיטה השנייה הייתה בחינת דגם הביטוי של עצים נורמאליים ועצים בעלי פנוטיפ ננסי, באמצעות ריצוף בהספק גבוה בשיטות DGE ו-mRNA-seq.

### 5.1. דגם המתילציה

הירידה ההדרגתית בחומרת הפנוטיפ של 'חנטה לקויה' ו'ריבוי שחלות' עם הזמן (Cohen *et al.*, 2004) מצביעה על האפשרות של מנגנון אפיגנטי היוצר את הטיפוסים החריגים. כמו כן עבודות קודמות הצביעו על קשר אפשרי בין מתילציה DNA במספר מקטעי DNA ובין דגם הביטוי של הגן DEF לבין פנוטיפים חריגים של תמר. בעבודה זאת בחנו את דגם המתילציה בארבעה מקטעי DNA בהם נמצאה שונות באתר *MspI* בין עצים נורמאליים לחריגים בעלי מופע ננסי ובעלי מופע של 'חנטה לקויה' ו'ריבוי שחלות' (גורביץ, 2004, עבודת גמר). כמו כן נבדק דגם המתילציה של הגן DEF בעצים מהזן 'ברה' בעלי 'חנטה לקויה' ו'ריבוי שחלות' בהם נמצאו בעבר הבדלים בדגם הביטוי (שוחט, 2007, עבודת גמר). דגם המתילציה נבחן באמצעות Bisulfite sequencing. בקרת ביטוי באמצעות מתילציה יכולה להתרחש ב-DNA או בחלבונים. עבודה זאת עסקה במתילציה של DNA.

#### 5.1.1. שונות בדגם המתילציה של מקטעי DNA שבודדו באמצעות MS-AFLP

אנליזת MS-AFLP נעשתה על ידי ס. גורביץ (גורביץ, 2004, עבודת גמר) לאפיון דגם המתילציה בתמרים מזן 'ברה' ו'מג'הול'. האנליזה התבססה על אנזימי רסטריקציה *MspI* ו-*HpaII* המזהים את אותו אתר חיתוך (CCGG), אך שונים ברגישותם למתילציה.

במקטע אחד (Bis6) נמצא הבדל ברמת המתילציה באתר *MspI* בין עצי WT ל-DW בדומה לתוצאות אנליזת ה-MS-AFLP. כמו כן בהשוואה של דגם המתילציה בכל אתרי CG במקטע, נראה שמקטע זה בעצי WT היה ממותל יותר מאשר בעצי ה-Off-type (96% לעומת 54.7%, ראה טבלה 9). גם במקטע Bis5 נמצא הבדל ברמת המתילציה כאשר עצי WT היו ממותלים ב-29.5% מאתרי CG לעומת עצי Off-type שבהם האתרים היו ממותלים רק ב-5.7% אולם באנליזת Bisulfite sequencing אתר ה-*MspI* עצמו לא נמצא ממותל.

בשלושה מהמקטעים (Bis1, Bis2, ו-Bis5) אתר *MspI* לא נמצא ממותל בעבודה זאת בניגוד לממצא מאנליזת ה-MS-AFLP. ממצא זה יכול לנבוע מכמה סיבות:

1. ריצוף המקטעים מאנליזת MS-AFLP לא היה תמיד באיכות מספקת. לעתים התקבל יותר מרצף אחד לכל פרגמנט. יתכן והרצף שבו בחנתי את דגם המתילציה אינו הרצף שבו נמצא הבדל באתר ה-*MspI* על ידי ס. גורביץ.

2. בחלק מהמקרים אומנם בודד רצף DNA אחד, אבל הרצף היה הומולוגי ליותר מרצף אחד בטיוטת הגנום של התמר. באחד המקרים, למרות שנמצאו שני רצפים אפשריים למקטע מסוים מאנליזת MS-AFLP, בחנתי את המתילציה רק באחד מהמקטעים (למקטע השני לא היה ניתן לבנות תחלים שיכללו את אתר *MspI* הרצוי, כיוון שטיוטת גנום התמר עדיין איננה מלאה).
3. אנליזת ה-Bisulfite sequencing בוצעה על שלושה עצים מכל פנוטיפ, יתכן שבפרטים אלה אתר ה-*MspI* הנבדק לא היה ממותל. המדגם הנבחן היה מאוד קטן ולא תמיד נבדקו אותם העצים שבהם נמצאה בעבר השונות באנליזת MS-AFLP. יתכן שהגדלת המדגם הייתה מביאה לתוצאות שונות.
4. מובן מאליו שאי אפשר לשלול שגיאות מסוגים שונים שיכולות להיווצר בכל אחת משיטות הבדיקה. לסיכום, בשניים מתוך ארבעת המקטעים שנבדקו נמצאה שונות בדגם המתילציה בין עצים נורמאליים לחריגים, רק באחד מהם הממצא היה תואם לאנליזת ה-MS-AFLP.

### 5.1.2. דגם ביטוי המתילציה של הגן DEF

נבדק דגם המתילציה של הגן DEF. גן זה נבחר בעקבות העבודה של מ.שוחט, בה נמצא שביטוי הגן DEF בעצים החריגים בעלי הפנוטיפים החמורים 'חנטה לקויה' ו'ריבוי שחלות' היה נמוך פי שלושה לעומת העצים הנורמאליים (שוחט, 2007, עבודת גמר). הגן DEF מכיל *homeobox* מסוג B החיוני להתמיינות עלי הכותרת והאבקנים בפרח. בדקל השמן נמצאו הבדלים ברמת הביטוי בין פרטים נורמאליים לפרטים חריגים בעלי פנוטיפ 'Mantled', של מספר גנים מסוג MADS box שאחד מהם היה הגן DEF (Adam *et al.*, 2007). בנוסף, בחינה ראשונית של המתילציה על ידי שימוש באנזימים *HpaII* ו-*MspI* שנעשתה על ידי מ.שוחט, הצביעה על מתילציה ברמה נמוכה בעצים חריגים אבל על היעדר מתילציה בעצים נורמאליים (שוחט, 2007, עבודת גמר).

לרמת המתילציה יש השפעה על התפתחות הפרח ועל פוריותו. שינויים ברמת המתילציה נמצאו כגורמים לעיוותים בפרחים, לתוספת אברים בפרח (אבקנים או שחלות), ולעקרות (Richardson, 1997; Kaeppler, 2000). בעבודות קודמות מצאו פולימורפיזם ברמת המתילציה של ה-DNA בצמחים שיוצרו בתרבית רקמה לעומת צמחים שרובו בשיטות קונבנציונאליות (Peraza-Echeverria *et al.*, 2001; Jaligot *et al.*, 2000). בדקל השמן נמצאה ירידה במתילציה בשיעור של 0.5-2.5% בהשוואת בפרטים עם פנוטיפ 'Mantled' לפרטים נורמאליים (Jaligot *et al.*, 2000).

החלק הנבדק הכיל את אזור ה-5' UTR ושלושת האקסונים הראשונים והיה באורך 1350 בסיסים. דגם המתילציה נבחן בשלושה עצים בעלי פנוטיפ נורמאלי לעומת שלושה עצים בעלי מופע של 'חנטה לקויה' ו'ריבוי שחלות'. הרקמות הנבדקות היו עלים ותפרחות. לא נמצאה מתילציה משמעותית במקטע זה בשתי הרקמות הנבדקות ולא נמצאו הבדלים משמעותיים בדגם המתילציה בין עצים נורמאליים לחריגים. אתרי ה-CCGG שנבחנו על ידי מאור שוחט ונמצאו ממותלים ברמה נמוכה בעצים חריגים לא נמצאו ממותלים כלל, לא ברקמת העלה (שנבחנו ע"י מ.שוחט) ולא בתפרחות. הבחינה של רמת המתילציה שנעשתה בעבר הייתה בחינה עקיפה המבוססת על חיתוך בשני אנזימי רסטריקציה בעלי רגישות שונה למתילציה ולאחריה

הגברה ב-PCR. בשיטה זאת עשוי להיות מצב בו מעט DNA שלא יחתך יעבור הגברה. לכן, יתכן שהתוצאות הנ"ל נבעו מפעילות לא מלאה של אנזימי הרסטריקציה, ולא כתוצאה ממתילציה ספציפית באותם אתרים. ראוי לציין שדגם המתילציה נבחן רק בחלק מהגן ולא בכולו, ועל כן יתכן שקיימת מתילציה באזורים אחרים של הגן שלא נבחנו על ידי. כמו כן יתכן שגנים שנמצאים upstream לגן DEF הם שמבוקרים על ידי מתילציה. בארבידופסיס נמצא שהגן AP3 (אורתולוג של DEF) מבוקר על ידי הגנים FLY ו-AP1. FLY נקשר לפרומוטור של AP3 ובכך מאפשר את הביטוי שלו. AP1 חיוני להצטברות RNA של UFO (unusual floral organs) במהלך השלב המאוחר של התפתחות הפרח והוא חיוני להתמיינות עלי הכותרת והאבקנים בפרח. הסתבר שגם במוטציה של loss-of-function של FLY ו-UFO זוהתה רמה נמוכה של AP3 RNA בקבוצה קטנה של תאים. כלומר יתכן שקיים מסלול נוסף של אקטיבציה של AP3 (Jack, 2004).

בעצי תמר חריגים שרובו בתרבית רקמה התגלתה שונות רבה בחומרת הפנוטיפ בין מטעים שונים ובין פרטים שונים באותן חלקות בני אותו גיל. השונות אינה רק ברמת 'החנטה הלקויה' אלא גם באחוז 'ריבוי השחלות'. במרבית העצים החריגים התגלה שיפור של רמות החנטה, וירידה ברמת הפירות המשולשים והפירות מרובי השחלות עם הזמן. תופעת 'ריבוי השחלות' נעלמת באופן הדרגתי לפני תופעת 'החנטה הלקויה' (Cohen et al., 2004). אם המנגנון האחראי על יצירת פנוטיפים אלה הוא אכן מתילציה, אזי עיתוי לקיחת הדוגמאות ובחינת המתילציה הוא מאוד משמעותי. כגן מבקר פריחה, DEF אמור להתבטא רק בשלבים מסוימים בהתפתחות הפרח, ורק ברקמות מאוד ספציפיות בו. המתילציה עשויה להשתנות במהלך ההתפתחות ויתכן שהדוגמאות בהם בחנתי מתילציה בגן לא נלקחו בעיתוי הנכון ולכן ולא נמצאה שונות בדגם המתילציה בין הפרטים החריגים לנורמאליים. בעבודה זאת דגם המתילציה של הגן DEF נבדק בעלים בוגרים בפרחים שלמים מתפרחות בוגרות שכבר סיימו את השלב ההתפתחותי שלהם. יתכן שבשלב זה הגן כבר איננו ממותל. כמו כן יתכן שהמתילציה קיימת אך היא לא נמצאת בכל רקמות והתאים אלא רק באחוז מסוים מהם. חשוב להדגיש ששיטת ה-Bisulfite sequencing שבה בחנתי מתילציה, אינה מתאימה לאיתור שיעור נמוך של תאים ממותלים.

עבודות קודמות מצאו פולימורפיזם ברמת המתילציה של ה-DNA בצמחים שיוצרו בתרבית רקמה לעומת צמחים שרובו בשיטות קונבנציונאליות (Peraza-Echeverria et al., 2001; Jaligot et al., 2000). בעבודה קודמת (גורביץ, 2004) בתמרים שבחנה שונות בדגם המתילציה בין עצים נורמאליים לכאלו שרובו בתרביות רקמה נמצא שונות רבה בדגם המתילציה בין עצים נורמאליים לעצים ננסים מהזן 'מג'הול'. התוצאות הצביעו על האפשרות שבמהלך הריבוי של הזן 'מג'הול' בתרביות רקמה, מצטברת שונות בדגם המתילציה. לעומת זאת לא נמצאו הבדלים משמעותיים בדגם המתילציה בין עצים נורמאליים לבעלי לפנוטיפ 'החנטה הלקויה' של הזן 'ברה' (גורביץ, 2004, עבודת גמר). הבדלים במתילציה של ה-DNA נמצאו גם בין עצי דקל השמן עם פנוטיפ 'Mantled' לעצים נורמאליים (Matthes et al., 2001; Jaligot et al., 2000, 2004). בבנות נמצאה מתילציה בעצי בננה ננסים במקטע מסוים. בעצי בננה נורמאליים מקטע זה לא נמצא ממותל (Oh et al., 2007). בעבודה זאת ההבדלים בדגם המתילציה שנמצאו היו מקטעים (Bis5 ו-Bis6) שמקורם



בזן 'מג'הול'. הבדלים אלה מראים שככל הנראה קיים קשר בין שינויים בדפוסי המתילציה בגנים ובאזורים מסוימים בגנום ליצירת פנוטיפים חריגים בתרביות רקמה.

## 5.2. אפיון דגם ביטוי באמצעות Solexa/Illumina

נעשה שימוש בשתי שיטות גנומית לצורך אפיון דגם הביטוי של גנים רבים בעצי תמר מזן 'מג'הול' שרובו באמצעות חוטרים או באמצעות תרביות רקמה. דגם ביטוי גנים נבדק באמצעות אנליזת ריצוף בנפח גבוה בשיטות DGE ו-mRNA-seq בטכנולוגיה של Solexa. באנליזת Digital Gene Expression מתקבלים תגים קצרים (כ-21 ו-27 בסיסים) מהקצה ה-3' של התעתיק. באנליזת mRNA-seq התקבלו רצפים של 33 בסיסים שעוברים הרכבה (assembly) ליצירת תעתיקים שלמים. היתרונות העיקריים של שימוש בטכנולוגיית Solexa/Illumina הוא שהאנליזה אפשרית גם באורגניזמים שאין עליהם ידע גנומי רב (Collins *et al.*, 2008). קיימים מספר הבדלים בין DGE ל-mRNA-seq: היתרון בשיטת DGE הוא הרגישות הרבה שלה. מכיוון שהיא מבוססת על תגים קצרים מהם לא צריך לבנות קונטיגים, כל תג נספר בפני עצמו ומייצג מולקולת RNA אחת שרוצפה. כך מתקבל פרופיל ביטוי שכולל בתוכו גם גנים מאוד נדירים. לעומת זאת מכיוון ש-DGE מתבסס על שימוש באנזימי רסטריקציה, רצפי התגים בשיטה זאת מייצגים את הקצה ה-3' של התעתיק. התגים עשויים לייצג אזורים לא מתורגמים של הגן (אזורי ה-3' UTR). כאשר אין מידע גנומי אמין על אזורים אלה, זה מקשה על מציאת הגן אליו שייך התג כי האזורים שאינם מתורגמים אינם שמורים ברמה גבוהה. בנוסף, חלק ממאגרי המידע לא תמיד מכילים מידע על הקצה ה-3' וכיוון שאנליזת DGE מבוססת על הנוכחות של אתר רסטריקציה של האנזים המשמש לעיכול ה-cDNA הדו-גדילי, חלק מהתעתיקים יכולים להיות חסרי אתר זה. תעתיקים אלה לא נכללו באנליזה. כדי להתגבר באופן חלקי על בעיה זאת, נבחנו ריצופים בנפח גבוה לאחר עיכול בחמישה אנזימים שונים, אולם החיבור וההשוואה בין התגים שנוצרו בשיטות השונות אינו אפשרי ללא מידע גנומי מקיף. חסרון נוסף הוא קבלת תגים מאוד קצרים ולכן תג מסוים יכול לייצג מספר גנים שונים. בהעדר רצף מלא של הגנום חסרון זה מקשה מאוד על מציאת הגנים אליהם שייך התג, במיוחד עם מאפשרים בחיפוש גם שינויים בבסיס אחד או שניים מהתג הקצר. שיטת mRNA-seq מבוססת על שבירת mRNA באמצעים מכאניים, ריצופו וביצוע assembly לקונטיגים. בשיטה זאת מתקבלים רצפים המייצגים את כלל ה-mRNA ברקמה. בעבודה זאת נעשה ריצוף של ה-cDNA משני קצוות הפרגמנט (paired end). פרוטוקול זה מאפשר קבלת שני רצפים לכל "מושבת DNA" ולכן מידע רב יותר על כל מקטע מרוצף. בגישה זאת ניתן גם להתייחס לריצוף משני הצדדים כאימות, ולבחון רק רצפים מ"מושבות DNA" בהם הרצפים משויכים לאותם קונטיגים בגנום על פי הרצפים משני הצדדים. לשיטת mRNA-seq יש מספר חסרונות והראשון הוא שנדרש ידע מוקדם על רצף הגנום לצורך ביצוע assembly של המקטעים. בנוסף, גם בשיטה זאת הרצפים המתקבלים הם יחסית קצרים (למרות היותם ארוכים יחסית לרצפים המתקבלים ב-DGE) ולכן חלק מרצפים אלה יכולים להיות משויכים למספר אזורים בגנום. באנליזה שלנו, התעלמנו מרצפים שלא ניתן היה לשייך לגן בודד בלבד. חשוב לציין שהטכנולוגיות

לריצוף בנפח גבוה מתפתחות במהירות רבה, וכבר במהלך עבודה זאת ניתן היה לרצף בטכנולוגיה זאת מקטעים ארוכים יותר (עד ל-120 בסיסים מכל צד).

### 5.3. השוואת האנליזות השונות ששימשו לאפיון דגם הביטוי

ריצוף בנפח גבוה נעשה בשתי שיטות: אנליזת DGE בוצעה על רקמת "לב" תמר של עץ אחד נורמאלי ואחד ננסי ואנליזת mRNA-seq בוצעה על רקמת "לב" תמר של שני עצים נורמאליים ו-15 עצים בעלי מופע ננסי. בכל אחת מהאנליזות חושב היחס בין רמת הביטוי של עצי WT ועצי DW בכל אחד מהתגים/גנים. נעשתה השוואה בין כל אחת מהשיטות ל-Q-RT-PCR. באנליזת ה-DGE נעשתה השוואה בין עץ נורמאלי אחד לחריג אחד. באנליזת ה-mRNA-seq נעשו שתי השוואות: הראשונה היא בין תוצאות mRNA-seq ל-Q-RT-PCR של עץ WT אחד מול תערובת של 15 עצי DW והשנייה בין תוצאות mRNA-seq ל-Q-RT-PCR של שני עצי WT מול שש עצי DW (לא בתערובת אלא כפרטים בודדים). השוואה נוספת נעשתה בין אנליזת ה-DGE ל-mRNA-seq. מקטעים שאותרו ב-DGE הווארכו ב-RACE 5' ונבחנו ב-Q-RT-PCR. מקטעים אלו אותרו ברצפים הצפויים להיות רצפים מקודדים בטיוטת הגנום, ונמצא דגם הביטוי שלהם. לסיכום סה"כ נעשו שלוש השוואות הבאות: השוואה בין אנליזת DGE ל-Q-RT-PCR, השוואה בין mRNA-seq ל-Q-RT-PCR והשוואה בין DGE ל-mRNA-seq.

#### 5.3.1. השוואה בין תוצאות ה-DGE לתוצאות ה-Q-RT-PCR

לא נמצא דמיון בין תוצאות ה-DGE לתוצאות ה-Q-RT-PCR. נבחנו 12 תגים שריצפם הווארך על ידי RACE 5'. נבחרו תגים בעלי יחסי ביטוי קיצוניים בין העץ הנורמאלי לעץ בעל פנוטיפ ננסי, אף אחד מהרצפים הללו לא נתן באנליזת ה-Q-RT-PCR יחס דומה (בין עץ נורמאלי לננס) לזה שהתקבל באנליזת ה-DGE. ברוב הרצפים התקבל יחס ביטוי קרוב לאחד. רק ברצף אחד התקבל יחס של 0.33 (הגן ב-DW מתבטא פי שלוש יותר מאשר ב-WT) בין הפרטים הנורמאליים לחריגים. לפי תוצאות ה-DGE גן זה ב-DW מתבטא פי 8 יותר מאשר ב-WT. קיימות מספר סיבות אפשריות לכך שלא הצלחנו לחזור על תוצאות ה-DGE:

1. יתכן שהרצפים שהווארכו באמצעות RACE 5' לא ייצגו את התג המקורי. ניתן היה ליצור טרנסקריפטום עדכני ל"לבבות" תמר למשל על ידי שימוש בטכנולוגיה של 454 בה הרצפים הם באורך של מאות בסיסים, ובכך יתכן שניתן היה לזהות את מרבית התגים ללא צורך בביצוע RACE 5'.
2. כיוון שהתגים של ה-DGE הם קצרים מאוד יתכן שתג מסוים מייצג יותר מתעתיק RNA אחד, ורק אחד מהם (ולאו דווקא העיקרי והמייצג), הווארך בריאקציה ה-RACE 5'. לכן היחס ברמת הביטוי המחושב בין הפרטים מייצג תעתיקים שונים. זה נכון במיוחד בגלל מיקומם של חלק ניכר מהתגים באזורים שאינם מתורגמים (3' UTR) ושומרים פחות בגנום. יתכן כמובן גם שבריאקציה ה-RACE 5' הווארכו תגים שהכילו שינוי בבסיס בודד או יותר ברצף (mismatch) ולכן קרוב לודאי גם קידדו לגנים אחרים.

### 5.3.2. השוואה בין תוצאות ה-mRNA-seq לתוצאות ה-Q-RT-PCR

נעשתה השוואה בין תוצאות ה-mRNA-seq לבין התוצאות שהתקבלו ב-Q-RT-PCR. אנליזת mRNA-seq בוצעה על תערובת של שני עצי WT ו-15 עצים ננסיים. אנליזת ה-Q-RT-PCR בוצעה על 24 גנים (טבלה 18) בעץ WT אחד (עקב חוסר ברקמת "לב" מאחר העצים) מול תערובת של 15 עצי תמר ננסיים. גן אחד בו לא היה הבדל ברמת הביטוי בין עצי WT ל-DW שימש כביקורת לאנליזת ה-Q-RT-PCR. באנליזת ה-Q-RT-PCR נתקבלה רמת ביטוי שונה בין ה-WT ל-DW ב-21 גנים מתוך ה-23 (בהם היה הבדל ברמת הביטוי באנליזת ה-mRNA-seq). ככל שיחס זה ברמת הביטוי היה גדול יותר באנליזת mRNA-seq כך היה גם באנליזת ה-Q-RT-PCR. באנליזה זאת לא התקבלו הבדלים משמעותיים בין ה-WT לתערובת DW בשני גנים: PDK\_20s1530771g001 ו-PDK\_20s1377181g001.

למרות ערכי הקורלציה הגבוהים בין אנליזת mRNA-seq ל-Q-RT-PCR, לא קיבלנו את אותם היחסים בדיוק. בדרך כלל ההבדלים ברמות הביטוי בין שני הפנוטיפים היו קטנים יותר באנליזת ה-Q-RT-PCR בהשוואה לתוצאות אנליזת ה-mRNA-seq. לכך יכולות להיות כמה סיבות:

1. שיטת ה-Q-RT-PCR מוגבלת עקב רקע שנוצר בגלל ההגברה של הגנים ב-PCR והשימוש בסמן פלואורסנטי. הרקע יכול לנבוע מהגברה לא שווה של מולקולות DNA שונות, הגברה לא ספציפית או של קטעי DNA שמקורם בזיהומים שונים ואי דיוקים הנובעים מהנפחים הזעירים הנלקחים לריאקציה.
2. אנליזת ה-Q-RT-PCR התבצעה על תערובת DW מול פרט WT אחד ולא מול תערובת של שני עצים נורמאליים כמו באנליזת mRNA-seq, זאת עקב חוסר ברקמת "לב" של אחד העצים. הבדל זה יכול היה להשפיע על היחסים שהתקבלו באנליזה. מכיוון שאנליזת ה-mRNA-seq מבוססת על ריצוף ישיר של מולקולות ה-cDNA, סביר שהיא אמינה ומדויקת יותר, אולם, גם בה יכולות להיווצר שגיאות.

נבחרו ארבעה גנים בעלי הבדלים גדולים מאוד ברמות ביטוי בין ה-WT ל-DW על פי אנליזת mRNA-seq ובהם נבדקה רמת הביטוי של עצים בודדים ב-Q-RT-PCR (שני עצי WT וששה עצי DW). גם באנליזה זאת קיבלנו הבדלים גדולים ברמת הביטוי בין עצי WT ל-DW בדומה לתוצאות mRNA-seq. היה דמיון רב יותר בין אנליזת mRNA-seq לאנליזת ה-Q-RT-PCR על פרטים בודדים מאשר בין mRNA-seq לאנליזת ה-Q-RT-PCR על תערובת. יתכן שהסיבה לכך היא שאנליזת ה-Q-RT-PCR התבצעה על תבנית של עץ WT בודד ואילו אנליזת ה-mRNA-seq נעשתה על שני עצי WT.

באנליזה שהתבצעה על עצים בודדים נמצאה לעתים שונות בין העצים הננסיים השונים (טבלה 19). בגן PDK\_20s1392771g001 לא הייתה שונות בין רוב העצים הננסיים הנבדקים אבל בפרט אחד היה ביטוי שונה מאוד של הגן (ביטוי נמוך יותר פי 110). בחלק מהמקרים נמצאה גם שונות בין שני העצים הנורמאליים. בגן PDK\_20s1685261g001 השונות בין העצים הנורמאליים הייתה מאוד גבוהה (כמעט פי 13). בגלל השונות הגבוהה ברמת הביטוי בין העצים הנורמאליים קשה להתייחס לשונות בין העצים הננסיים, שבהם השונות היא עד פי 32 בין העצים. בשאר הגנים, השונות בין שני העצים הנורמאליים הייתה נמוכה ובגן PDK\_20s1955911g001 כמעט ולא נמצאה שונות בין העצים הננסיים. כדי להקטין את השונות בין העצים היה צריך לבצע את האנליזה על מספר גבוה יותר של עצי WT.

אנליזת mRNA-seq נעשתה על תערובות RNA. לעבודה עם תערובות יש יתרונות וחסרונות שונים. היתרון העיקרי הוא שבהרצה אחת ניתן לבחון הרבה פרטים במקביל ובעלות יחסית נמוכה. גישה זאת נכונה כשאנו מחפשים הבדלים משמעותיים ברמות הביטוי בין שתי התערובות הנבדקות המייצגות את הפרטים השונים בכל אחת מהן. החיסרון העיקרי הוא שלא ניתן לבחון את השונות בין הפרטים. חסרון נוסף הוא אפשרות שגיאה כתוצאה מחוסר דיוק, שכן כמויות ה-RNA צריכות להיות מדויקות מאוד בכדי ליצור תערובת מייצגת. אנליזת Q-RT-PCR על פרטים בודדים התבצעה רק פעם אחת עם חזרות פנימיות בלבד (שלוש חזרות על כל מיהול). גם אנליזת הריצוף בנפח גבוה התבצעה בחזרה אחת בלבד. כדי לקבל תוצאות אמינות ומדויקות יותר יש צורך לבצע חזרות נוספות ובלתי תלויות.

### 5.3.3. השוואה בין תוצאות ה-DGE לתוצאות ה-mRNA-seq

נעשה אפיון של 12 תוצרי RACE 5' (שאותרו ב-DGE ונבדקו ב-Q-RT-PCR) מול הרצפים הצפויים להיות רצפים מקודדים בטיוטת גנום התמר (predicted coding sequence). לשלושה מתוך ה-12 לא נמצאה הומולוגיה לרצפים המקודדים הצפויים. תשעה רצפים אחרים היו הומולוגים לרצפי mRNA הצפויים ושלושה מתוכם היו הומולוגים לשני רצפי mRNA בעלי אנוטיציה דומה. תשעת הרצפים שהיו הומולוגים לרצפי mRNA הצפויים נמצאו באנליזת SOAP ונבחן בהם היחס בין WT ל-DW (טבלה 21). נמצא שבשבעה מתוכם היחס שהתקבל בשני האנליזות (Q-RT-PCR ו-mRNA-seq) היה קרוב ליחס של 1:1. עבור רצף אחד התקבל יחס של פי ארבעה לפי mRNA-seq ופי שלושה לפי Q-RT-PCR. ברצף נוסף נתקבל יחס של 0.1 באנליזת ה-mRNA-seq, אך לא נמצא הבדל משמעותי ביחסים בין WT ל-DW באנליזת ה-Q-RT-PCR. לסיכום היחסים שהתקבלו באנליזת ה-mRNA-seq עבור הרצפים שהוארכו ב-RACE 5' היו דומים ליחסים שהתקבלו באנליזת ה-Q-RT-PCR (מקדם קורלציה 0.826 ומובהקות 0.006). ובכך יש אימות לתוצאות אלה. כמו כן ראוי לציין את אימות התוצאות מאנליזת ה-mRNA-seq באמצעות Q-RT-PCR (ראה להלן) עבור 24 גנים. יתכן שהבעיה העיקרית באנליזת ה-DGE וההארכה ב-RACE 5' הינה שמרבית הרצפים שהוארכו בעזרת RACE 5' לא היו אותם הרצפים בהם היו הבדלים משמעותיים בתגים הקצרים ב-DGE. ידוע שבהארכת מקטעי cDNA באמצעות RACE 5' מתקבלים פעמים רבות תוצרים לא נכונים. כדי לצמצם תופעה זו, ולקבל בעיקר הארכה ספציפית של רצף ה-cDNA, משתמשים פעמים רבות בתחלים ארוכים מאוד או בהארכה בשני שלבים – ובשני תחלים באזורים שונים של הרצף המקורי (nested primers), אולם גישות אלו לא היו אפשריות בהתחשב באורך התגים הקצר שהיה ברשותנו. לכן, בתוצרי ה-RACE 5' שהתקבלו לא נמצאו הבדלים בין ה-WT ל-DW באנליזת ה-DGE.

### 5.3.4. גנים בהם נמצאו הבדלים ברמת הביטוי בין פרטים נורמאליים לחריגים

באנליזת mRNA-seq בלטו מספר גנים שנבדלו ברמת הביטוי בין עצים נורמאליים לננסיים והקשורים למספר תהליכים ביולוגיים. הגנים הנ"ל מעורבים בתהליכי ביוסינטזה של אוקסין, תהליכי חמצון-חיזור, זירחון

של חומצות אמינו, הובלה טרנס-ממברנאלית, בקרה על שעתוק ועוד. כן אותרו גנים הקשורים לצימוח, יצירה ובקרה של פעילות הורמונאלית (ג'יברלין, ברסינוסטרואידים ואוקסין), גנים לננסות המעורבים בהתארכות תאים בצמחים כמו אקספנסין וגנים הקשורים למתילציה (methyltransferase).

לאנליזת ה-Q-RT-PCR נבחרו כמה גנים בהם הייתה ירידה ברמת הביטוי בעצי DW לעומת עצי WT וביניהם גנים במסלול הביוסינטזה של ההורמון ג'יברלין, הגן לאקספנסין, פקטור השעתוק WRKY, ומספר גנים בהם הייתה עליה ברמת הביטוי בעצי DW לעומת עצי WT באנליזת Q-RT-PCR ומשתתפים ביצירת אוקסין, BR1 (brassinosteroid insensitive 1) והגן diminuto (חלבון המעורב בחישה של BR).

הורמונים צמחיים רבים כמו אוקסין, ברסינוסטרואידים, ג'יברלין ואתילן ממלאים תפקיד חשוב בבקרה של גדילת הצמח ומתווכים את התגובה הביוטית והאביוטית לעקות (Santner *et al.*, 2009). ג'יברלין הוא הורמון הגדילה העיקרי המווסת את התארכות הגזע, נביטה ופריחה. עבודות רבות מתארות קשר בין ג'יברלין לבין ננסות. באורז ובעגבניות ננסיות נמצאה רמת ג'יברלין נמוכה בהרבה מזאת שבפרטים הנורמאליים (Kobayashi *et al.*, 1989; Fray *et al.*, 1995). באנליזת mRNA-seq ו-Q-RT-PCR הראיתי שרמת הביטוי של PDK\_20s1306264g001 (בעל אנוטציה ל-20 gibberellin) גבוהה יותר בעצים נורמאליים פי 27 באנליזת mRNA-seq ופי כ-6.7 באנליזת Q-RT-PCR מאשר בעצים בעלי פנוטיפ ננסי (טבלה 18). בארבידופסיס זוהתה מוטציה הגורמת לננסות כתוצאה מכניסה של טרנספוזון לגן shi. המוטציה מקנה פנוטיפ האופייני למוטנטים הפגועים בתגובה לג'יברלינים (Fridborg *et al.*, 1999). כן ידוע על שימוש במעכבי צמיחה המעכבים יצירה של ג'יברלינים וגורמת לעיכוב צימוח וננסות. צמחי *Pharbitis nil* המטופלים במרסן צמיחה מסוג CCC (trimethylammonium chloride) היו ננסיים ומעוכבים ביצירת הפרחים (Zeevaart, 1966). באורז נמדדה השפעה של prohexadone-calcium על רמת הג'יברלינים. נמצא שאורך הצמח והעלים קטנו בכ-78% ו-36% בהתאמה, ורמת הג'יברלינים GA<sub>19</sub> ו-GA<sub>20</sub> ירדה בצמחים המטופלים (Nakayama *et al.*, 1992). בתמרים נמצא שפקלובטרזול ויוניקונזול, מעכבי סינטזת ג'יברלין, גרמו לצמצום בולט של צימוח הגזע וירידה מסוימת בגודלם העלים (אלוני וחוב' 2009; Aloni *et al.*, 2010). בעבודה אחרת נמצא שבאורז חצי-ננסי (semidwarf) האלל sd-1 (הקשור לפנוטיפ) גורם לחוסר בג'יברלין אקטיבי הנובע מ-GA 20 oxidase לקוי. במצב תקין GA 20 oxidase מבצע אוקסידציה ל-GA53 ומתקבלים GA20 ו-GA1 פעילים. באורז חצי-ננסי נמצא שרמת הסובסטרט ההתחלתית (GA53) של GA 20 oxidase מצטבר ואילו רמת התוצרים שלו (GA20 ו-GA1) נמוכה יותר (Spielmeyer *et al.*, 2002). אוקסין משפיע על קצב גדילת איברי הצמח השונים. בדקל השמן זוהו שני גנים EgM39A ו-EgIAA1 שרמת הביטוי שלהם הייתה גבוהה יותר בעצים בעלי הפנוטיפ 'Mantled' שרובו בתרבות רקמה. EgIAA1 הוא חלבון המעורב בתגובה מוקדמת לאוקסין ו-EgM39A מקודד לחלבון לא ידוע המראה דמיון ברצף ל-asparagines synthetases (Morcillo *et al.*, 2006). בתירס ננסי היה יצור מופחת של אוקסין ביחס לצמחים נורמאליים (Overbeek, 1935). צמחי ארבידופסיס בעלי המוטציה SAX1 הגורמת לרגישות לאוקסין הראו פנוטיפ ננסי, עיכוב בהתפתחות וירידה בפוריות. (Ephritikhine *et al.*, 1999). אני בחנתי ארבעה גנים בעלי אנוטציה לפקטורי שעתוק של אוקסין. בכל ארבעתם מצאתי עליה ברמת הביטוי בעצי תמר ננסיים לעומת העצים הנורמאליים באנליזת mRNA-seq ו-Q-RT-PCR (טבלה 18).

ברסינוסטרואידים (BS) הם הורמונים צמחיים המעודדים את התפתחות הגבעול, הבשלת הזרעים וויסות גנים רבים הקשורים להתפתחות הצמח. מוטנטים בגנים לסינתזת או לחישת ברסינוסטרואידים הינם בדרך כלל ירוקים מאוד וננסיים. באורז אופיינה מוטציה רצסיבית לננסות הנגרמת כתוצאה ממוטציה באנזים BR-6-oxidase המשתתף בביוסינתזה של BS. מתן BS חיצוני שיחזר בצורה טובה את הפנוטיפ הנורמאלי (Mori *et al.*, 2002). אני בחנתי באמצעות Q-RT-PCR שלושה גנים שהראו הומולוגיה ל-BRI1 (brassinosteroid insensitive 1). באנליזת mRNA-seq נמצאה עליה ברמת הביטוי של שלושת הגנים האלה בעצי DW (טבלה 18). BRI1 הינו רצפטור שקושר ברסינוסטרואידים. עליה ברמת הביטוי שלו צפויה לגרום לעליה ברגישות לברסינוסטרואידים ואולי דווקא לצימוח מוגבר של הצמח. עם זאת נמצא כי צמחי ארבידופסיס ננסיים חסרי האלל BRI1 (Brassinosteroid Insensitive 1) לא מגיבים ל-BS מבחון. בצמחי ארבידופסיס חסרי BRI1 קיימת הצטברות של כמויות גדולות של BS, castasterone ו-typhasterol פי 57, 128 ו-33 בהתאמה בהשוואה לצמחי WT. בצמחי ארבידופסיס בעלי ביטוי חלש של האלל BRI1 יש הצטברות של הורמונים אלה אבל בכמויות נמוכות יותר (Noguchi *et al.*, 1999). יתכן שהצטברות זאת של BRI1 בעצי תמר ננסיים יכולה להביא לחוסר יכולת לנצל BS או לחוסר יכולת לווסת את הביוסינתזה של BS. יתכן ששלושת הרצפים ההומולוגיים ל-BRI1 שבחנתי, מקודדים לאזורים שונים של אותו גן. בידוד הגן המלא בתמר ובחינתו המלאה תשפר את ההבנה שלנו לגבי הגורם המשפיע על הצטברות BS בעצים ננסיים.

גן נוסף הקשור לביוסינתזה של ברסינוסטרואידים הוא הגן *diminuto/dwarf1*. בצמחי ארבידופסיס ננסיים (*dim mutant*) יש ירידה משמעותית בפוריות. צמחים אלה לא מבטאים mRNA של הגן *diminuto/dwarf1* (Takahashi *et al.*, 1995). נמצא שגן זה מקודד לחלבון המעורב בסינתזת סטרואידים, הוא מפחית את האיזומריזציה של הקשר  $\Delta^{24(28)}$ , קשר זה הוא הראשון שעובר איזומריזציה בתהליך יצירת סטרול מסוג brassinolide, ההכרחי לביוסינתזה של סטרולים בצמחים. הוספה של brassinolide (סוג של ברסינוסטרואיד) מחזירה את הפנוטיפ הנורמאלי (Klahre *et al.*, 1998). מוטציית חסר בגן *dwarf1* של העגבנייה, גורמת לננסות חמורה ואי אספקה של ברסינוסטרואידים (Nomura *et al.*, 2005). באנליזת mRNA-seq ו-Q-RT-PCR נמצא שהגן *diminuto/dwarf1* עולה בעצי ה-DW לעומת עצי ה-WT פי 117 ו-25 בהתאמה. תוצאה זאת מנוגדת לממצאים שונים המראים שגן זה יורד בצמחים ננסיים (Klahre *et al.*, 1998; Nomura *et al.*, 2005). נדרש אפיון של רצף כל הגן והחלבון כדי להשוות את הממצאים עם הידוע בספרות. יתכן שלמרות שרמת ה-mRNA של הגן עולה בעצים ננסיים לא נוצר החלבון או נוצר חלבון לא פעיל ומכאן פנוטיפ הננסות. באורז זוהה מוטנט הפגוע במסלול היצירה של הברסינוסטרואיד *brd2*. למוטנט זה יש חסר של בסיס אחד באזור המקודד של הגן *diminuto/dwarf1* האורתולוגי מארבידופסיס. צמחי אורז מוטנטים ב-*brd2* בעלי פנוטיפ *semi-dwarf* בשלב הווגטיבי, הם בעלי ליקויים חמורים בהתארכות הפרקים ובהתפתחות המכבד (*panicle*) והזרעים (Hong *et al.*, 2005). מעניין לציין שראינו עליה בביטוי של שני גנים הקשורים לביוסינתזה ולחישת של ברסינוסטרואידים, *dwarf1* ו-BRI1 בעצי DW. שני גנים אלה קשורים אחד לשני בכך ש-*dwarf1* קשור לסינתזה של ברסינוסטרואידים ואילו BRI1 הוא הרצפטור התאי שחש בברסינוסטרואידים. העלייה ברמת הביטוי שלהם יכולה להצביע על

עליה בחשיבות הברסינואידים בעצים הננסיים. יתכן שזוהי תגובה של העץ לפנוטיפ הננסות (שנגרמת אולי בגלל שינויים ופגיעה במסלולים ביוסנתטיים להורמונים אחרים כמו ג'יברלינים או אוקסינים) על ידי הגברת הרגישות של מסלול ה-BS ו/או מסלולים נוספים. אם היפותזה זו נכונה, התמר מנסה "לפצות" על פנוטיפ הננסות על ידי הגברת הגן dwarf1 ליצירה מוגברת של ברסינוסטרוואידים והגברה של BRI1 כדי שיהיה שיפור בחישה שלהם, אז שני גנים אלה אינם גורמים לפנוטיפ הננסות אלא דווקא מושפעים ממנו.

גן נוסף שנמצא להתבטא באופן שונה בין עצי WT לעצי DW הוא האקספנסין. אקספנסין הוא חלבון המצוי בדופן תאי הצמחי ולו תפקידים חשובים בגדילת התא, התרככות הפרי, תיפקוד המריסטמה ותהליכים התפתחותיים אחרים (Cosgrove, 2000). למרות זאת, מוטנטים באקספנסין אינם צמחים ננסיים. נמצא שלפחות חלק מהגנים לאקספנסין מבוקרים על ידי הורמונים שונים במהלך הגדילה (Jung *et al.*, 2010). נמצא, שמוטנטים של BS בעלי הפחתה בהתארכות התאים מראים גם רמת mRNA נמוכה יותר של אקספנסין ו-XET (xyloglucanase endotransglycosylase) (Kauschmann *et al.*, 1996). Cosgrove ו-יצרו בארבידופסיס קווי Antisense בהם הושקו הגן AtEXP10 המקודד לאקספנסין, ומצאו שצמחים אלה היו בעלי עלים קטנים יותר, ודופן התא שלהם הייתה בעל יכולת מתיחה קטנה יותר. הם הראו שלאקספנסין תפקידים התפתחותיים רבים שמבקרים את גודל האברים ואת המורפולוגיה (Cho and Cosgrove, 2000). בעבודה זאת ראינו ירידה ברמת הביטוי של אקספנסין, בעצי תמר ננסיים באנליזות mRNA-seq ו-Q-RT-PCR פי 14 ו-16.4 בהתאמה, מה שיכול גם כן להסביר את פנוטיפ הננסות. בעתיד כדאי יהיה לבחון את גודל התאים באזורים השונים של לב התמר (וברקמות אחרות) ולנסות לזהות האם הננסות נובעת גם מתאים קצרים יותר.

באנליזת mRNA-seq ו-Q-RT-PCR נמצא שרמת הביטוי של הגן PDK\_20s1368561g002 (בעל אנוטציה לפקטור שעתוק WRKY) ירדה פי 133 ופי 12 בהתאמה. גן זה שייך למשפחה גדולה של פקטורי שעתוק המעורבים בתהליכים פיזיולוגיים של גדילה, התפתחות, תגובה לעקות ביוטיות, אביוטיות ולהזדקנות (senescence) (Ulker and Somssich, 2004). לפקטורי שעתוק יש תפקיד חשוב בתהליך השעתוק. עיכוב פקטור שעתוק מסוים יכול להוביל לאי יצירה או ליצירה ברמה נמוכה של גנים מסוימים. מחקרים מראים שחלק מחלבוני WRKY הם בקרי מפתח בשלבי האמבריוגנזה (Luo; Lagace and Matton, 2004; Luo *et al.*, 2005). רמות נמוכות של פקטור שעתוק זה עשויות להיות הגורם לפנוטיפ הננסות. באנליזת mRNA-seq נמצאו עוד 8 גנים המקודדים לפקטורי שעתוק WRKY, חמישה מהם בעלי רמת ביטוי גבוהה יותר בעצי WT ושלושה מהם בעלי רמת ביטוי גבוהה יותר בעצי DW (לא הוצג בפרק התוצאות). ניתן להסביר תוצאות אלה בכך שפקטורי שעתוק אלה משפיעים על מגוון רחב של גנים אשר רמת הביטוי של חלקם עולה ושל אחרים יורדת. מעניין יהיה לבודד ולבחון את מגוון הגנים הללו ואת תפקידם.

Miura וחוב' (2009) מצאו שהגן dwarf1 באורז נמצא תחת בקרה אפיגנטית ומשפיע על גובה האורז. שינויים במתילציה של אתר הפרומוטור של גן זה, גורמים להשתקה של הגן וליצירת פנוטיפ הננסות (Miura *et al.*, 2009). נמצא שטיפול בצמח Bryonia ב-5-azacyd (מעכב של DNA methyltransferase) הביא לפנוטיפ של ננסות (Galaud *et al.*, 1993). בדקל השמן נמצא שרמת השעתוק של שנים מתוך שלושה גנים המקודדים ל-DNA מתילטרנספראז (DNMT) גבוהה יותר בעצים בעלי פנוטיפ 'Mantled' המרובים

בתרביות רקמה לעומת עצים נורמאליים (Rival *et al.*, 2008). מכאן ניתן להסיק שלמתילציית ה-DNA יש השפעה על פנוטיפ הננסות בצמחים שונים. בעבודה הנוכחית לא נבחנו רמות DNA methyltransferase בעצים נורמאליים וחריגים באמצעות Q-RT-PCR. באנליזת mRNA-seq נמצאה עליה ברמות הביטוי של שלושה גנים בעלי אנוטיציה ל-DNA methyltransferase פי 2.3, 5.7 ו-10.8 בעצים ננסים לעומת עצים נורמאליים. כמו כן נמצאה עליה ברמות הביטוי של שלושה גנים בעלי אנוטיציה ל-Histone methyltransferase פי 2.7, 5.6 ו-5.7 בעצים ננסים לעומת עצים נורמאליים (תוצאות אלה לא אומתו ב-Q-RT-PCR). דרושה בחינה של רמת הביטוי שלהם באנליזת Q-RT-PCR לאימות של התוצאות.

## 5.4. סיכום וכיוונים עתידיים

מטרת עבודה זאת הייתה לאפיין תהליכי יצירת שונות סומקלונלית על ידי זיהוי הבדלים בין עצי תמר חריגים מתרבית רקמה לעצים נורמאליים. ההבדלים בפנוטיפים בין עצים נורמאליים לעצים חריגים יכולים לנבוע מכמה סיבות: הבדלים ברצף ה-DNA, הבדלים בדגם הביטוי של גנים מסוימים, הבדלים בדגם המתילציה אשר גורמים הבדלים בדגמי הביטוי והבדלים בהרכב וכמות אנזימים ו/או החלבונים מסוימים. בעבודה הנוכחית נבחנו הבדלים בדגם הביטוי ובדגם המתילציה בין עצים נורמאליים לחריגים. יש לציין שדגם המתילציה יכול להשפיע על דגם הביטוי.

דגם המתילציה נבדק בארבעה מקטעי DNA ובגן DEF בהם נמצאו הבדלים בעבודות קודמות (גורביץ, 2004, עבודת גמר; שוחט, 2007, עבודת גמר). רק באחד המקטעים (Bis6) נמצאה התאמה לתוצאות הקודמות. במקטע נוסף (Bis5) נמצא הבדל בדגם המתילציה בין פרטים נורמאליים לחריגים (אך רק במושבה אחת נמצא שהאתר *MspI* ממותל כפי שנמצא באנליזת MS-AFLP). יתכן שההבדל במתילציה במקטעי DNA אלה קשור למנגנון יצירת השונות הסומקלונלית בתרביות רקמה אך בעבודה זאת לא נמצאה לכך הוכחה. שני מקטעי DNA אחרים (Bis1 ו-Bis2) והגן DEF לא הראו הבדלים בדגם המתילציה בין פרטים נורמאליים לחריגים. תוצאה זאת יכולה להתקבל עקב סיבות טכניות (ראה לעיל) או שמקטעים אלה וגן זה אינם ממותלים ברקמות שנבדקו. יתכן שבעתיד כדאי לבחון מתילציה באזורים רחבים יותר של ה-DNA, בגנים נוספים ובאזורי הפרומוטור שלהם שיכולים להיות קשורים למנגנון יצירת שונות סומקלונלית כמו למשל גנים שמבקרים את הגן DEF. דרושה גם בחינה של דגם המתילציה בשלבי התפתחות שונים. בנוסף דרושה הרחבה של מספר העצים והרקמות הנבדקות כדי שניתן יהיה לקשור בין דגמי המתילציה לפנוטיפ. כיוון נוסף הוא בחינת דגם המתילציה בגנום השלם של מספר עצים נורמאליים ומספר עצים חריגים ברקמות שונות בדומה לעבודה שנעשתה בארבידוספסיס (Zhang *et al.*, 2006).

בבחינת דגמי הביטוי באמצעות אנליזת mRNA-seq נמצאו (לאחר סינונים שונים – ראה תוצאות) כ-2,000 גנים בעלי הבדלים ברמת הביטוי בין עצים נורמאליים לחריגים. נבחרו 23 גנים שרמת הביטוי שלהם נבחנה גם באנליזת Q-RT-PCR. 21 מתוך 23 גנים אלה הראו הבדלים ברמת הביטוי גם זאת. למרות שנמצאו גנים רבים בעלי הבדלים ברמת הביטוי בין עצים נורמאליים לחריגים ממצאים אלה אינם מצביעים באופן ישיר על המנגנון של יצירת שונות סומקלונלית בתרביות רקמה. ההבדלים ברמת הביטוי בין עצים נורמאליים לחריגים



יכולים לנבוע כתוצאה מתנאי הריבוי בתרבית. יש לציין שהאנליזה בתמר מבוצעת על רקמת "לב" התמר ולצורך הוצאת הרקמה צריך להקריב עץ שלם שעלותו גבוהה. ההרחבה במספר העצים ובמספר הרקמות ובשלב גידול שונים תוכל לצמצם את מספר הגנים לאלה שנבדלים רק בין (between) הפנוטיפים ולא בתוכם (within). כמו כן אנליזת mRNA-seq על עצים בודדים תוכל לצמצם את השונות בין עצים נורמאליים לחריגים. כיוון נוסף הוא השוואה של אנליזות ריצוף בנפח גבוה בין מיני צמחים שונים המרובים בתרבות רקמה בעלי פנוטיפים דומים כמו ננסות בתמרים ובנות או 'חנטה לקויה' ו'ריבוי שחלות' בתמרים ובדקל השמן. אפשר להניח שבעתיד הרחוק יותר ניתן יהיה לעשות ריצוף של הגנום השלם של מספר פרטים נורמאליים ומספר פרטים חריגים ולהשוות ביניהם. אנליזה כזאת דורשת טכנולוגית ריצוף זולה ומדויקת שאינה זמינה היום.

#### **5.4.1. אימות תוצאות וגישות לחקר המנגנון האחראי ליצירת השונות הסומקלונלית**

כדי לאפיין את המנגנון הגורם ליצירת שונות סומקלונלית בתמרים נדרשות אנליזות נוספות שיכולות לאמת את הקשר בין ממצאים מולקולאריים ושונות פנוטיפית. מציאת גנים בעלי הבדלים בדגם המתלציה או דגם הביטוי בין עצים נורמאליים לחריגים אינה מספקת כדי לטעון שגנים אלה אחראים על יצירת הפנוטיפים החריגים (casual agents) כי הבדלים אלה יכולים להימצא בגנים שהם התוצאה לפנוטיפ ולא הסיבה שלו. רק מספר גנים מועט יהיו הגורמים הישירים לתהליך, ומרבית הגנים שימצאו שונים יהיו גנים שמושפעים מהשינוי שחל.

מקטעי DNA שנמצאו שונים בדגם המתלציה שלהם בין עצים נורמאליים לננסיים וכן גנים בעלי הבדלים ברמת הביטוי בין עצים נורמאליים לחריגים שנמצאו באנליזת mRNA-seq צריכים לעבור מבחנים נוספים. תחילה דרוש צמצום של רשימת המקטעים בעלי דגם ביטוי שונה בין עצים נורמאליים לחריגים. הקריטריונים לצמצום עשויים לכלול: הומולוגיה לגנים באורגניזמים אחרים בעלי פנוטיפ הקשור לננסות (או לפנוטיפ נבדק אחר), גנים בעלי הבדלים קיצוניים ברמת הביטוי שלהם בין עצים נורמאליים לחריגים וגנים הקשורים לתהליכי צימוח. כמו כן ניתן לבחון גנים המבקרים את הגנים בהם נמצאו הבדלים ברמת הביטוי. בשלב הבא יש לבצע אימות במספר מישורים: תחילה יש לאמת ולקבוע את מהימנותם וזאת על ידי שימוש בביקורות מתאימות ובמספר חזרות ראוי. שנית יש לבחון האם הגן הנבחן קשור לפנוטיפ או שההבדל שנמצא הינו אקראי. ולבסוף יש לנסות ולקבוע האם הגן הנבדק הוא זה הגורם לפנוטיפ או שהוא תוצאה שלו.

הדרך להתמודדות עם האפשרות של תוצאות אקראיות היא מובהקות סטטיסטית. לצורך כך יש להרחיב את מספר העצים הנורמאליים והחריגים עליהם נעשית הבדיקה (בעבודה זאת אפיון דגם המתלציה ודגם הביטוי). כמו כן יש לבצע את האנליזות על מספר גדול של גנים ובמספר חזרות ראוי לקבלת מובהקות סטטיסטית. כדי לבחון האם הגן הנבחן יוצר את הפנוטיפ החריג או שהוא התוצאה שלו ניתן לבצע שינוי של הגנוטיפ על ידי טרנסגניים. מכיוון שלא פותחו שיטות טרנספורמציה בתמר לא ניתן לבחון במערכת זאת השתקה או ביטוי יתר של גנים. ניתן לעשות זאת במערכות הטרולוגיות בצמחי מודל כמו ארבידופסיס או אורז (שהינו צמח חד

פסיגי ולכן דומה יותר לתמר). דרך נוספת להתמודד עם בעיית הטרנספורמציה בתמרים היא הגדלה של מספר העצים הנבדקים, שכן גם בבני אדם לא ניתן לבצע טרנספורמציה של גנים ובכל זאת בהרבה מחלות ידועות מוטציות ספציפיות בגנים ספציפיים הגורמים למחלה. בבני אדם הגנים האחראים על מחלות זוהו על ידי בחינה של מספר גדול של חולים לעומת בריאים. ניתן ליישם גישה זאת גם בתמרים ולבחון מספר ראוי של עצים חריגים בעלי אותו הפנוטיפ מול מספר גדול של עצים נורמאליים. גישה זאת קשה יותר לביצוע כאשר הפנוטיפ הנחקר הוא תוצאה של אינטראקציה בין מספר גנים ואללים.

## 5.5. חשיבות חקלאית

אפיון גנים המצויים באסוציאציה לפנוטיפ הננסות ופנוטיפ 'חנטה לקויה' ו'ריבוי שחלות' עשויים לשמש בעתיד לזיהוי מוקדם של הפנוטיפים החריגים. זיהוי מוקדם יכול למנוע שתילת עצים במטע בהם הפנוטיפ מתגלה רק לאחר מספר שנות גידול, ובכך למנוע אובדן השקעות רבות. בנוסף, הבנת הבקרה הגנטית של תהליכי הצימוח חשובה להבנת הבקרה על גובה העץ במטע ותוכל לסייע להגדרת התנאים האופטימאליים של תרבית הרקמה לצורכי ריבוי שתילים ולפיתוח זנים בעלי גובה אופטימאלי. בזנים שונים של בנות זוהו טיפוסים ננסיים ו"חצי ננסיים", וחלקם הפכו בזכות תכונות אלה לזנים מסחריים. גם בדקלי השמן זוהו זנים חצי-ננסיים המשמשים מקור ליצירת זנים חדשים נמוכים יותר (Corley *et al.*, 1971; Hardon *et al.*, 1972; Luyindula *et al.*, 2005). הבנת המנגנון המולקולארי המגביל את הצימוח הווגטיבי יכול לתרום לענף התמרים ולענפים נוספים.

התופעה של שונות סומקלונלית מתרחשת במיני רבים נוספים כמו דקל השמן, בנות אורז ועוד. למידה והבנת של המנגנון הגורם לשונות סומקלונלית יכולה למנוע יצירת הטיפוסים החריגים או לחילופין לעזור בזיהויים וסילוקם לפני הנטיעה.

## 6. רשימת ספרות

- אלוני, ד.ד., פומרנץ, ר., חזון, ה., קרפ, ח., סנדלהם, ד., גלבע, א., כהן, י. (2009) עיכוב היתמרות בתמר. עלון הנוטע, 63, 22-27.
- ברנשטיין, צ. (2004). התמר. הוצאת הדיקלאים בע"מ ומועצת הפירות.
- גורביץ, ו. (2004). אפיון מולקולארי של השונות הגנטית בעצי תמר שמקורם בזנים שונים ובשיטות ריבוי שונות. עבודת גמר לשם תואר "מוסמך למדעי החקלאות", האוניברסיטה העברית בירושלים, הפקולטה למדעי החקלאות המזון ואיכות הסביבה.
- ישראלי, י. (2004) ריבוי תמרים בתרבית רקמה. מתוך: התמר (ברנשטיין, צ. (עורך)). הוצאת הדיקלאים בע"מ ומועצת הפירות.
- כהן, י., קורצ'נסקי, ר., טריפלר, א., זיו, ג. וכהן, ר. (2003) בעיות חנטה בעצי תמר מזן 'ברהי' שמקורם בתרבית רקמה. עלון הנוטע 57: 275-279.
- סטולר, ש. (1977). גידול התמר בארץ ישראל. הוצאת הקיבוץ המאוחד.
- שוחט, מ. (2007). אפיון גנים המבקרים את התפתחות אברי הפרה בתמר (*Phoenix dactylifera* L.) והשוואתם בין עצים נורמאליים ועצים בעלי פנוטיפים חריגים. עבודת גמר לשם תואר "מוסמך למדעי החקלאות", האוניברסיטה העברית בירושלים, הפקולטה למדעי החקלאות המזון ואיכות הסביבה.
- Adam, H., Jouannic, S., Morcillo, F., Orioux, Y., Duval, Y. and Tregear, J.W. (2007). Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *Journal of Experimental Botany* 58, 1245-1259.
- Adams, R.L.P. (1995). Eukaryotic DNA methyltransferases - structure and function. *BioEssays* 17, 139-45.
- Aitchitt, M., Ainsworth, C.C. and Thangavelu, M. (1993). A rapid and efficient method for the extraction of total DNA from mature leaves of the date palm (*Phoenix dactylifera* L.). *Plant Molecular Biology Reporter* 11, 317-319.
- Ali, T.A., Jubrail, J.M. and Jasim, A.M. (2007). The Use of RAPDs for the Detection of Genetic Stability of Regenerated Plantlets of Barhee Palm in Iraq. *Acta Horticulturae* 736, 127-134.
- Al Kaabi, H.H., Zaid, A. and Ainsworth, C. (2007a). Plant-off-types in tissue culture-derived date palm (*Phoenix dactylifera* L.) plants. *Acta Horticulturae* 736, 267-281.
- Al Kaabi, H.H., Zaid, A., Shephard, H. and Ainsworth, C. (2007b). AFLP Variation in Tissue Culture-Derived Date Palm (*Phoenix dactylifera* L.) Plants. *Acta Horticulturae* 736, 135-160.
- Aloni, D.D., Hazon, H., Edom, U., Karp, C., Pumeranc, R., Sandelham, D. and Cohen, Y. (2010). Effects of growth retardants on vegetative growth of date palms. *Acta Horticulturae*, in press.
- Al-Wasel, A.S. (2001). Field performance of somaclonal variants of tissue culture derived date palm (*Phonix dactylifera* L.). *Plant Tissue Culture* 11, 97-105.
- Bairu, M.W., Fennell, C.W. and Van Staden, J. (2006). The effect of plant growth regulators on somaclonal variation in Cavendish banana (*Musa* AAA cv. 'Zelig'). *Scientia Horticulturae* 108, 347-351.
- Barrow, S.C. (1998). A monograph of *Phoenix* L. (*Palmae: Coryphoideae*). *Kew Bulletin* 53, 513-5.

- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57, 289-300.
- Bentley, D.R. (2006). Whole-genome re-sequencing. *Current Opinion in Genetics & Development* 16, 545-552.
- Bhaskaran, S. and Smith, R. (1992). Somatic embryogenesis from shoot tip and immature inflorescence of *Phoenix dactylifera* cv. Barhee. *Plant Cell Reports* 12, 22-25.
- Bolstad, B.M., Irizarry, R.A., Astrand, M. and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinformatics* 19, 185-193.
- Bommert, P., Satoh-Nagasawa, N., Jackson, D. and Hirano, H.Y. (2005). Genetics and evolution of inflorescence and flower development in grasses. *Plant Cell Physiology* 46, 69-78.
- Busch, W. and Lohmann, J.U. (2007). Profiling a plant: expression analysis in Arabidopsis. *Current Opinion in Plant Biology* 10, 136-141.
- Chang, S., Puryear, J. and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11, 113-116.
- Chao, C.T. and Krueger, R.R. (2007). The date palm (*Phoenix dactylifera* L.): Overview of biology, uses, and cultivation. *Hortscience* 42, 1077-1082.
- Cho, H.T. and Cosgrove, D.J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 97, 9783-9788.
- Cohen, Y., Korchinsky, R. and Tripler, E. (2004). Flower abnormalities cause abnormal fruit setting in tissue culture - propagated date palms (*Phoenix dactylifera* L.) *Journal of horticultural science & biotechnology* 79, 1007-1013.
- Cohen, Y., Gurevich, V., Korchinsky, R., Shochat, M., Sekar, M. and Lavi, U. (2007). Molecular and phenotypic characterization of somaclonal variation in date palm off-types originated from tissue culture. *Acta Horticulturae* 738, 417-423.
- Collins, L.J., Biggs, P.J., Voelckel, C. and Joly, S. (2008). An approach to transcriptome analysis of non-model organisms using short-read sequences. *Genome Informatics* 21, 3-14.
- Cosgrove, D.J. (2000). Loosening of plant cell walls by expansins. *Nature* 407, 321-326.
- Corley, R.H.V., Hardon, J.J. and Tan, G.Y. (1971). Analysis of growth of the oil palm (*Elaeis guineensis* Jacq.) I. Estimation of growth parameters and application in breeding. *Euphytica* 20, 307-315.
- Corley, R.H.V., Lee, C.H., Law, I.M. and Wong, C.Y. (1986). Abnormal flower development in oil palm clones. *Planter* 62, 233-240.
- Corniquel, B. and Mercier, L. (1994). Date palm (*Phoenix dactylifera* L.) cultivar identification by RFLP and RAPD. *Plant Science* 101, 163-172.
- Cullis, C.A., Cullis, M.A. and Abdullah, M.O. (2007). Development of markers for the mantled phenotype in oil palm. In: PIPOC 2007. Proceedings of the Agriculture, Biotechnology and Sustainability Conference, Kuala Lumpur, Malaysia, pp. 299-312.

- Damasco, O.P., Graham, G.C., Henry, R.J., Adkins, S.W., Smith, M.K. and Godwin, I.D. (1996). Random amplified polymorphic DNA (RAPD) detection of dwarf off-types in micropropagated Cavendish (*Musa* spp. AAA) bananas. *Plant Cell Reports* 16, 118-123.
- Damasco, O.P., Adkins, S.W., Godwin, I.D. and Smith, M.K. (1998). Use of A SCAR-based marker for the early detection of dwarf off-types in micropropagated Cavendish bananas. *Acta Horticulturae* 461, 157-164.
- Djerbi, M. (2000). Abnormal fruting of the date palm derived from tissue culture. In: *Proceedings of the Date Palm International Symposium, held in Windhoek, Namibia, 22-25 February 2000*. Windhoek, p. 73, 257-260.
- Eeuwens, C.J., Lord, S., Donough, C.R., Rao, V., Vallejo, G. and Nelson, S. (2002). Effects of tissue culture conditions during embryoid multiplication on the incidence of 'Mantled' flowering in clonally propagated oil palm. *Plant Cell, Tissue and Organ Culture* 70, 301-323.
- Ehrlich, M. and Ehrlich, K.C. (1993). Effect of DNA methylation on the binding of vertebrate and plant proteins to DNA. *Experientia Supplementum* 64, 145-168.
- Ephritikhine, G., Fellner, M., Vannini, C., Lapous, D. and Barbier-Brygoo, H. (1999). The sax1 dwarf mutant of *Arabidopsis thaliana* shows altered sensitivity of growth responses to abscisic acid, auxin, gibberellins and ethylene and is partially rescued by exogenous brassinosteroid. *The Plant Journal* 18, 303-314.
- Eshraghi, P., Zarghami, R. and Ofoghi, H. (2005). Genetic stability of micropropagated plantlets in Date Palm. *Scientific information database* 16, 311-315.
- Evans, D.A., Sharp, W.R. and Medina-Filho, H.P. (1984). Somaclonal and gametoclonal variation. *American Journal of Botany* 71, 759-774.
- Feng, L., Liu, H., Liu, Y., Lu, Z., Guo, G., Guo, S., Zheng, H., Gao, Y., Cheng, S., Wang, J., Zhang, K. and Zhang, Y. (2010). Power of Deep Sequencing and Agilent Microarray for Gene Expression Profiling Study. *Molecular Biotechnology* 45, 101-110.
- Finnegan, E.J., Genger, R.K., Peacock, W.J. and Dennis, E.S. (1998). DNA methylation in plants. *Annual Reviews of Plant Physiology and Plant Molecular Biology* 49, 223-247.
- Finnegan, E., Peacock, W. and Dennis, E. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proceedings of the National Academy of Sciences USA* 93, 8449-8454.
- Fridborg, I., Kuusk, S., Moritz, T. and Sundberg, E. (1999). The *Arabidopsis* dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. *Plant Cell* 11, 1019-1031.
- Fray, R.G., Wallace, A., Fraser, P.D., Valero, D., Hedden, P., Bramley, P.M. and Grierson, D. (1995). Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. *The Plant Journal* 8, 693-701.
- Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences USA* 89, 1827-1831.

- Galaud, J.P., Gaspar, T. and Boyer, N. (1993). Effect of anti-DNA methylation drugs on growth, level of methylated DNA, peroxidase activity and ethylene production of *Bryonia dioica* internodes. *Physiologia Plantarum* 87, 528-534.
- Gruenbaum, Y., Naveh-Many, T., Cedar, H. and Razin, A. (1981). Sequence specificity of methylation in higher plant DNA. *Nature* 292: 860-862.
- Gupta, M., Chyi, Y.S., Romeo-Severson, J. and Owen, J.L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and Applied Genetics* 89, 998-1006.
- Gurevich, V., Lavi, U. and Cohen, Y. (2005). Genetic variation in date palms from offshoots and tissue culture. *Journal of the American Society for Horticultural Science* 130, 46-53.
- Hafiz, I.A., Anjum, M.A., Grewal A.G. and Chaudhary, G.A. (2001) DNA methylation - an essential mechanism in plant molecular biology. *Acta Physiologiae Plantarum* 23, 491-499.
- Hall, H. (2007). Advanced sequencing technologies and their wider impact in microbiology. *Journal of Experimental Biology* 209, 1518-1525.
- Hamama, L., Cornee, N., Leclerc, V., Marionnet, F., Javouhey, M. and Letouze, R. (2003). Date palm (*Phoenix dactylifera*) offshoot identification by PCR-ISSR markers. *Acta Horticulturae* 616, 453-457.
- Hardon, J.J., Corley, R.H.V. and Ool, S.C. (1972). Analysis of growth in the oil palm II. Estimation of genetic variances of growth parameters and yield of fruit bunches. *Euphytica*. 21, 257-264.
- Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M. (1996). Real Time RT-PCR. *Genome Research* 6, 986-994.
- Higuchi, R., Fockler, C., Dollinger, G. and Watson, R. (1993). Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (NY)* 11, 1026-1030.
- Hong, Z., Ueguchi-Tanaka, M., Fujioka, S., Takatsuto, S., Yoshida, S., Hasegawa, Y., Ashikari, M., Kitano, H. and Matsuoka, M. (2005). The Rice brassinosteroid-deficient dwarf2 mutant, defective in the rice homolog of Arabidopsis DIMINUTO/DWARF1, is rescued by the endogenously accumulated alternative bioactive brassinosteroid, dolichosterone. *Plant Cell* 17, 2243–2254.
- Jack, T. (2004). Molecular and genetic mechanisms of floral control. *Plant Cell* 16, S1-S17.
- Jaligot, E., Rival, A., Beule, T., Dussert, S. and Verdeil, J.L. (2000). Somaclonal variation in oil palm (*Elaeis guineensis* Jacq): the DNA methylation hypothesis. *Plant Cell Reports* 19, 684-690.
- Jaligot, E., Beule, T., Baurens, F.S., Billotte, N. and Rival, A. (2004). Search for methylation-sensitive amplification polymorphisms associated with the 'mantled' variant phenotype in oil palm (*Elaeis guineensis* Jacq.). *Genome* 47, 224-228.
- Johnson, C., Cullis, T.A., Cullis, M.A. and Cullis, C.A. (2009). DNA markers for variety identification in date palm (*Phoenix dactylifera* L.). *Journal of Horticultural Science & Biotechnology* 84, 591–594.
- Jost, J.P. and Saluz, H.P. (1993). DNA Methylation: Molecular Biology and Biological Significance. Basel: Springer-Verlag.

- Jung, J., O'Donoghue, E.M., Dijkwel, P.P. and Brummell, D.A. (2010). Expression of multiple expansin genes is associated with cell expansion in potato organs. *Plant Science* 179, 77-85.
- Kaeppler, S.M., Kaeppler, H.F. and Rhee, Y. (2000). Epigenetic aspects of somaclonal variation in plants. *Plant Molecular Biology* 43, 179-188.
- Kauschmann, A., Jessop, A., Koncz, C., Szekeres, M., Willmitzer, L. and Altmann, T. (1996). Genetic evidence for an essential role of brassinosteroids in plant development. *The Plant* 9, 701-713.
- Klahre, U., Noguchi, T., Fujioka, S., Takatsuto, S., Yokota, T., Nomura, T., Yoshida, S. and Chua, N.H. (1998). The Arabidopsis DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. *The Plant Cell* 10, 1677-1690.
- Kobayashi, M., Sakurai, A., Saka, H. and Takahashi, N. (1989). Quantitative Analysis of Endogenous Gibberellins in Normal and Dwarf Cultivars of Rice. *Plant and Cell Physiology* 30, 963-969.
- Kunert, K.J., Baaziz, M. and Cullis, C.A. (2003). Techniques for determination of true-to-type date palm (*Phoenix dactylifera* L.) plants: a literature review. *Emirates Journal of Agricultural Sciences* 15, 1-16.
- Lagace, M. and Matton, D.P. (2004). Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta* 219, 185-189.
- Larkin, P. J. and Scowcroft, W.R. (1981). Somaclonal variation - a novel source of variability from cell culture for plant improvement. *Theoretical and Applied Genetics* 60, 197-214.
- Lewis, J. and Bird, A. (1991). DNA methylation and chromatin structure. *FEBS letters* 285, 155-159.
- Linacero, R., Alves, E.F. and Vazquez, A.M. (2000). Hot spots of DNA instability revealed through the study of somaclonal variation in rye. *Theoretical and Applied Genetics* 100, 506-511.
- Lisitsyn, N., Lisitsyn, N. and Wigler, M. (1993). Cloning the differences between two complex genomes. *Science* 259, 946-951.
- Luo, M., Dennis, E.S., Berger, F., Peacock, W.J. and Chaudhury, A. (2005). MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. *The Proceedings of the National Academy of Sciences USA*. 102, 17531-17536.
- Luyindula, N., Corley, R.H.V. and Mantantu, N. (2005). A comparison of the Deli Dumpy and Pobé dwarf short stemmed oil palms and their outcrossed progenies. *Journal of Oil Palm Research* 17, 152-159.
- Matthes, M., Singh, R., Cheah, S.C. and Karp, A. (2001). Variation in oil palm (*Elaeis guineensis* Jacq.) tissue culture-derived regenerants revealed by AFLPs with methylation-sensitive enzymes. *Theoretical and Applied Genetics* 102, 971-979.
- Marguerat, S., Wilhelm, B.T. and Bahler, J. (2008). Next-generation sequencing: applications beyond genomes. *Biochemical Society Transactions* 36, 1091-1096.

- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bembien, L.A., Berka, J., Braverman, M.S., Chen, Y.J. and Chen, Z. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.
- McCubbin, M.J., Van Staden, J. and Zaid, A. (2000). A Southern African survey conducted for off-types on date palms produced using somatic embryogenesis. *Proceedings of the Date Palm International Symposium, Windhoek, Namibia*, pp. 68-72.
- McCutcheon, J.P. and Moran, N.A. (2007). Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proceedings of the National Academy of Sciences* 104, 19392–19397.
- Messeguer, R., Ganal, M.W., Steffens, J.C. and Tanksley, S.D. (1991). Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. *Plant Molecular Biology* 16, 753-770.
- Mingliang, X., Xiangqian, L. and Schuyler, S.K. (2000). AFLP-Based Detection of DNA Methylation. *Plant Molecular Biology Reporter* 18, 361-368.
- Miura, K., Agetsuma, M., Kitano, H., Yoshimura, A., Matsuoka, M., Jacobsen, S.E. and Ashikari, M. (2009). A metastable DWARF1 epigenetic mutant affecting plant stature in rice. *The Proceedings of the National Academy of Sciences* 106, 11218–11223.
- Morcillo, F., Gagneur, C., Adam, H., Richaud, F., Singh, R., Cheah, S.C., Rival, A., Duval, Y. and Tregear, J.W. (2006). Somaclonal variation in micropropagated oil palm. Characterization of two novel genes with enhanced expression in epigenetically abnormal cell lines and in response to auxin. *Tree Physiology* 26, 585–594.
- Mori, M., Nomura, T., Ooka, H., Ishizaka, M., Yokota, T., Sugimoto, K., Okabe, K., Kajiwara, H., Satoh, K., Yamamoto, K., Hirochika, H. and Kikuchi, S. (2002). Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis. *Plant Physiology* 130, 1152–1161.
- Nagasawa, N., Miyoshi, M., Sano, Y., Satoh, H., Hirano, H.Y., Sakai, H. and Nagato, Y. (2003). SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development* 130, 705-718.
- Nakayama, I., Kobayashi, M., Kamiya, Y., Abe, H. and Sakurai, A. (1992). Effects of a Plant-Growth Regulator, Prohexadione-Calcium (BX-112), on the Endogenous Levels of Gibberellins in Rice. *Plant and Cell Physiology* 33, 59-62.
- Noguchi, T., Fujioka, S., Choe, S., Takatsuto, S., Yoshida, S., Yuan, H., Feldmann, K.A. and Tax, F.E. (1999). Brassinosteroid-insensitive dwarf mutants of Arabidopsis accumulate brassinosteroids. *Plant Physiology* 121, 743-52.
- Nomura, T., Kushiro, T., Yokota, T., Kamiya, Y., Bishop, G.J. and Yamaguchi, S. J. (2005). The last reaction producing brassinolide is catalyzed by cytochrome P-450s, CYP85A3 in tomato and CYP85A2 in Arabidopsis. *The Journal of Biological Chemistry* 280, 17873–17879.
- Oh, T.J., Cullis, M.A., Kunert, K., Engelborghs, I., Swennenc, R. and Cullis, C.A. (2007). Genomic changes associated with somaclonal variation in banana (*Musa* spp.). *Physiologia Plantarum* 129, 766–774.
- Overbeek, J.V. (1935) The growth hormone and the dwarf type of growth in corn. *Proceedings of the National Academy of Sciences*. 21, 292-299.



- Pelissier, T., Thalmeir, S., Kempe, D., Sanger, H.L. and Wassenger, M. (1999). Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. *Nucleic Acids Research* 27, 1625-1634.
- Peraza-Echeverria, S., Herrera-Valencia, V.A. and Kay, A. (2001). Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). *Plant Science* 161, 359-367.
- Phillips, R.L., Kaeppler, S.M. and Olhoft, P. (1994). Genetic instability of plant tissue cultures: breakdown of normal controls. *Proceedings of the National Academy of Sciences USA* 91, 5222-5226.
- Reuveni, O., Adato, Y. and Lilien-Kipnis, H. (1972). A study of new and rapid methods for the vegetative propagation of Date Palm. *Date growers' institute* 49, 17-24.
- Reuveni, O. and Israeli, Y. (1990). Measures to reduce somaclonal variation in *in vitro* propagated bananas. *Acta Horticulturae* 275, 307-313.
- Richards, E.J. (1997). DNA methylation and plant development. *Trends in Genetics* 13, 319-323.
- Rival, A., Beulé, T., Barre, P., Hamon, S., Duval, Y. and Noirot, M. (1997). Comparative flow cytometric estimation of nuclear DNA content in oil palm (*Elaeis guineensis* Jacq.) tissue cultures and seed-derived plants. *Plant Cell Rep* 16, 884-887.
- Rival, A., Bertrand, L., Beulé, T., Combes, C., Trouslot, P. and Lashermes, P. (1998a). Suitability of RAPD analysis for the detection of somaclonal variants in oil palm (*Elaeis guineensis* Jacq.). *Plant Breed* 117, 73-76.
- Rival, A., Jaligot, E., Beule, T. and Finnegan, E.J. (2008). Isolation and expression analysis of genes encoding MET, CMT, and DRM methyltransferases in oil palm (*Elaeis guineensis* Jacq.) in relation to the 'mantled' somaclonal variation. *Journal of Experimental Botany* 59, 3271-3281.
- Rival, A., Jaligot, E., Beulé, T., Tregear, J.W. and Finnegan, J. (2009). The oil palm "mantled" somaclonal variation: a model for epigenetic studies in higher plants. *Acta Horticulturae* 829, 177-182.
- Rival, A., Tregear, J., Verdeil, J.L., Richaud, F., Beule, T., Hartman, C., Rode, A. and Duval, U. (1998b). Molecular search for mRNA and genomic markers of the oil palm 'mantled' somaclonal variation. *Acta Horticulturae* 461, 165-171.
- Sahijram, L., Soneji, J.R. and Bollemma, K.T. (2003). Analyzing somaclonal variation in micropropagated banana (*Musa* spp.). *In Vitro Cellular and Development Biology* 39, 551-556.
- Salman, R.M., Al Jibouri, A.A.M., Al Quadhy, W.K. and Omar, M.S. (1988). Isozyme and chromosomal analyses of tissue culture derived date palms. *Date Palm Journal* 6, 401-411.
- Saker, M.M., Adawy, S.S., Mohamed, A.A. and El-Itriby, H.A. (2006). Monitoring of cultivar identity in tissue culture-derived date palms using RAPD and AFLP analysis. *Biologia Plantarum* 50, 198-204.
- Saker, M.M., Bekheet, S.A., Taha, H.S., Fahmy, A.S. and Moursy, H.A. (2000). Detection of somaclonal variations in tissue culture-derived date palm plants using isoenzyme analysis and RAPD fingerprints. *Biologia Plantarum* 43, 347-351.
- Santner, A., Calderon-Villalobos, L.I. and Estelle, M. (2009). Plant hormones are versatile chemical regulators of plant growth. *Nature Chemical Biology* 5, 301-307.

- Shimizu, T.S., Takahashi, K. and Tomita, M. (1997). CpG distribution patterns in methylated and non-methylated species. *Gene* 205, 103-107.
- Smith, M.K. and Drew, R.A. (1990). Growth and yield characteristics of dwarf off-types recovered from tissue-cultured bananas. *Australian Journal of Experimental Agriculture* 30, 575–578.
- Sokal, R.R. and Rohlf, F.J. (1995). *Biometry: The principle and practice of statistics in biological research*, Third Edition edn. Freeman, New York.
- Spielmeyer, W., Ellis, M.H. And Chandler, P.M. (2002). Semidwarf (sd-1), “green revolution” rice, contains defective gibberellins 20-oxidase gene. *Proceedings of the National Academy of Sciences* 99, 9043-9048.
- Takahashi, T., Gasch, A., Nishizawa, N. and Chua, N.H. (1995). The DIMINUTO gene of *Arabidopsis* is involved in regulating cell elongation. *Genes & Development* 9, 97–107.
- Tisserat, B. (1979). Propagation of date palm (*Phoenix dactylifera* L.) in vitro. *Journal of Experimental Botany* 30, 1275-1283.
- Ulker, B. and Somssich, I.E. (2004). WRKY transcription factors: from DNA binding towards biological function. *Current Opinion in Plant Biology* 7, 491-498.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T., Hornes, M. A. F., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407-4414.
- Wang, M.B. and Waterhouse, P.M. (2000). High-efficiency silencing of a l-glucuronidase gene in rice is correlated with repetitive transgene structure and is independent of DNA methylation. *Plant Molecular Biology* 43, 67-82.
- Wang, X., Elling, A., Li, X., Li, N., Peng, Z., He, G., Sun, H., Qi, Y., Liu, X.S. and Wang Denga, X. (2009). Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in Maize. *The Plant Cell* 21, 1053–1069.
- Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18, 7213–7218.
- Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J. and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453, 1239-1243.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6531–6535.
- Wu, J., Zhang, S., Meng, Q., Cao, H., Li, Z., Li, X., Shi, S., Kim, D., Bi, L., Turro, N. and Ju, J. (2007). 3'-O-modified nucleotides as reversible terminators for pyrosequencing. *Proceedings of the National Academy of Sciences USA* 104, 16462-16467.
- Zaid, A. (1999). Date palm cultivation. ISSN 0259-2517 FAO Plant production and protection paper number 156. FAO, Rome.
- Zeevaart, J.A.D. (1966) Reduction of the Gibberellin Content of Pharbitis Seeds by CCC and After-Effects in the Progeny. *Plant Physiology* 41, 856-862.

- Zehdi, S., Sakka, H., Ould, A.M.S., Rhouma, A., Marrakchi, M. and Trifi, M. (2005). Molecular polymorphism and genetic relationships in a Tunisian date-palm (*Phoenix dactylifera* L.) collection using ISSR amplification fingerprinting. *Plant Genetic Resources Newsletter* 144, 39-44.
- Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S., Chen, H., Henderson, I., Shinn, P., Pellegrini, M., Jacobsen, S. and Ecker, J. (2006). Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell* 126, 1189–1201.

## Summary

Date palm is a major tree crop in arid regions of the Middle East and North Africa. Date palms are traditionally propagated through offshoots, which are limited in number. The global need to expand date palm groves had led to the development of methods for tissue culture propagation. Tissue culture enables the generation of large number of plants identical to their “Mother” tree. However, a significant number of these plants have abnormal phenotypes. This phenomenon is named Somaclonal Variation, it occurs in many plants propagated through tissue culture and results in major economic losses.

The current study focuses on molecular characterization of two common off-type phenotypes in date palms: The ‘dwarf’ phenotype in the cultivar ‘Medjool’ and the ‘low fruit-setting’ and ‘multicarpel’ flowers in ‘Barhee’. We focused on characterization of patterns of DNA methylation and gene expression in both normal and off-type trees. Methylation patterns were analyzed by “Bisulfite Sequencing”, and gene-expression was assessed by high-throughput sequencing by the Solexa/Illumina technology and confirmed by Q-RT-PCR.

The methylation patterns were examined in several sites of genomic DNA: (1) Four genomic DNA fragments, where methylation patterns have been previously found to differ between normal and off-type trees by MS-AFLP analysis. (2) The DEFICIENS (DEF) gene which has been previously found to have a lower level of transcription in the ‘low fruit-setting’ and ‘multicarpel’ flowers compared to the normal trees.

In this study we found that two of the DNA fragments have a different methylation pattern in leaves of the normal and the dwarf tree. However, in the other two fragments, no differences were found in the methylation pattern between the normal and the ‘low fruit-setting’ and ‘Multicarpel’ flowers. The methylation level of the DEF gene was found to be very low. Moreover, no differences in the methylation pattern were found between the leaves and flowers of the normal and the ‘low fruit-setting’ and ‘multicarpel’ flowers.

The pattern of gene expression was characterized in “Hearts” of the normal and the dwarf ‘Medjool’ trees. The characterization was performed using two different protocols of high-throughput sequencing by the Solexa/Illumina technology: Digital Gene Expression (DGE) and mRNA-seq. The DGE analysis was carried out on one normal and one dwarf tree. This approach is based on the generation of several million short tags (21 or 27 base pairs) by the restriction enzymes: *Nla*III, *Hin*P1I, *Hpy*CH4IV, *Hpa*II, and *Taq*I, followed by amplification and sequencing. Tags were filtered to obtain expression ratios of at least three-fold between the normal and the off-type phenotypes. A total of 627, 155, 36, 48 and 172 tags having different

expression patterns in dwarf and normal trees were detected by the various enzymes, respectively. One hundred TAGs were selected to be elongated by 5' RACE. Twenty four long sequences were selected and analyzed by Q-RT-PCR. In most of the sequences, the Q-RT-PCR did not confirm the results of the DGE. A gene expression ratio of about one was obtained, in most of the Q-RT-PCR analyses.

In mRNA-seq, the RNA was fragmented and 33 base pairs were sequenced from both ends of each fragment (Paired-end analysis). The sequences were assembled based on the date palm genome draft. The mRNA-seq was performed on RNA mixtures of two normal trees and 15 dwarf trees. 13,647,010 sequences were generated in the RNA mixture of normal trees, while 14,443,018 sequences were generated in the mixture of the dwarf trees, having 450,351,330 and 476,619,594 base pairs, respectively. The sequences were mapped and assembled using the SOAP software to 17,032 predicted genes. Approximately 2,000 genes, having different expression patterns in dwarf and normal trees, were identified after filtering (see above). Twenty three genes were selected for Q-RT-PCR analysis. Twenty one of them differed in their pattern of gene-expression between normal and dwarf trees, confirming the mRNA-seq results. The correlation between the two analyses was statistically significant (correlation coefficient: 0.992). In addition, four of these genes were analyzed by Q-RT-PCR in six single dwarf trees and two normal trees. The results of the mRNA-seq were confirmed by Q-RT-PCR by a significant correlation coefficient (0.989).

Among the genes having different expression patterns, between normal and dwarf trees based on the mRNA-seq analysis, we have identified several interesting groups including: Genes related to growth, biosynthetic pathways, regulation of hormones and DNA methylation. Mutations in several of these genes have been identified as causing dwarf phenotypes in other plants. These include genes in biosynthesis and signaling of Gibberellin, Auxin and Brassinosteroid hormones, as well as genes for Expansin and the WRKY transcription factors.

In summary, several DNA fragments and a large number of genes, having different methylation and gene expression patterns between the normal and off-type phenotypes, were identified in the current study. Future gene analyses have the potential to contribute to the understanding of the mechanism that regulates the generation of off-type in date palms. Continuation of this study may improve the tissue culture protocols and lead to the development of new molecular tools for early identification of off-types. At the same time, it is important to emphasize the difficulties in the characterization of this mechanism. It is hopeful that future studies will be relevant to the identification of off-type phenotypes in various plant species propagated through tissue culture, as well as for understanding the mechanism of Somaclonal Variation.

Summary of the main findings:

1. A difference in the methylation pattern in two DNA fragments was found between a normal and a dwarf trees.
2. In a genomic analysis of the transcriptome, 2.000 genes were found to have a difference in the gene expression patterns between normal and dwarf trees. 21 of them were verified by Q-RT-PCR.

This work was carried out under the supervision of:

Dr. Cohen Yuval, Department of Fruit Trees Sciences, The Volcani Center.

Prof. Lavi Uri, Department of Fruit Trees Sciences, The Volcani Center.

Prof. Perl-Treves Rafael, Department of Life sciences, Bar-Ilan University.

**BAR-ILAN UNIVERSITY**

**Molecular analysis of somaclonal variants in date palm (*Phoenix dactylifera* L.)**

Anat Romanovsky

Submitted in partial fulfillment of the requirements for the Master's Degree  
The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University