

דוח מסכם לתכנית מחקר מספר 364-0191-8-00

## יישום סלקציה בעזרת הפלוטיפים בטיפוח בקר לחלב

### APPLICATION OF HAPLOTYPE ASSISTED SELECTION IN DAIRY CATTLE

מוגש להנהלת ענף בקר

ע"י

איל סרוסי	מכון בע"ח, מינהל המחקר החקלאי
מיכה רון	מכון בע"ח, מינהל המחקר החקלאי
יהודה ולר	מכון בע"ח, מינהל המחקר החקלאי

Eyal Seroussi Institute of Animal Science, ARO, P.O.B. 6, Bet-Dagan, 50250 Israel

E mail: seroussi@agri.huji.ac.il

Micha Ron Institute of Animal Science, ARO, P.O.B. 6, Bet-Dagan, 50250 Israel

E mail: micha@agri.huji.ac.il

Joel Weller Institute of Animal Science, ARO, P.O.B. 6, Bet-Dagan, 50250 Israel

E mail: weller@agri.huji.ac.il

נובמבר 2008

חשון תשס"ט

## הממצאים בדו"ח זה הינם תוצאות ניסויים.

הניסויים מהווים המלצות לחקלאים: **כן/לא** \* מחקר את המיותר

חתימת החוקר

\*

### רשימת פרסומים

Baruch, E., Weller, J.I., Cohen, M., Ron, M., Seroussi, E. (2006) Efficient Inference of Haplotypes from Genotypes on a Large Animal Pedigree. *Genetics* **172**, 1757-1765.

Seroussi, Y., Seroussi, E. (2007) *TraceHaplotype*: Using direct sequencing to determine the phase of an indel followed by biallelic SNPs. *BioTechniques* **43**, 452-456.

Seroussi, E. (2008) Concordance test emerges as a powerful tool for identifying quantitative trait nucleotides: lessons from BTA6 milk yield QTL. *Animal Genet.* in press.

## 2.1.1 תקציר

**הצגת הבעיה:** במהלך סריקת הגנום מופו אתרים לתכונות כלכליות כמותיות בבקר לחלב בישראל. זיהוי מדויק של הגנים האחראים ואיתור השונות הגנטית בהם הם משימות מורכבות ויקרות. ניתן לבצע סלקציה לגן כמותי בהעדר זיהוי של הגן עצמו באמצעות סלקציה להפלוטיפ, הכולל את האלל הרצוי. הפלוטיפ הוא רצף מסוים של אללים המאפיין קבוצת סמנים גנטיים קרובים על פני הכרומוזום התוחמים את האתר. סמנים מסוג של שינוי בבסיס יחיד (SNP) הם נפוצים בגנום, נוחים לזיהוי ולקביעת גנוטיפים ממוכנת. אנליזה של סמנים במבנה משפחתי מאפשרת קביעת הפלוטיפים באוכלוסיה הנבדקת. למרות ששיטות ההסקה הן סדרתיות ומובנות, הן מורכבות למדיי לביצוע אנושי בפרק זמן סביר עבור מספר סמנים רב באוכלוסיית הבקר.

**מטרות המחקר:** לפתח תוכנת מחשב לקביעת הפלוטיפים המותאמת לעדר הבקר ולבחור אלגוריתם שיסייע במיפוי של הגנים כמותיים על בסיס חוסר שיווי משקל באוכלוסיה בעזרת הפלוטיפים.

**שיטות ומהלך העבודה:** נאסף חומר גנטי מ-716 פרים עם אומדני הורשה שנקבעו על בסיס מבחן צאצאים. אוכלוסיית פרים זאת היא בעלת קשרי משפחה מסועפים המאפשרים מעקב אחר הפלוטיפים השכיחים בעדר הישראלי. גנוטיפים לסמנים גנטיים נקבעו באזורים הקריטיים בכרומוזום 6 ו-7. תוכנות נכתבו ב-C++ וב-PERL.

**תוצאות:** במסגרת מחקר זה, פורסם מאמר המתאר תוכנת מחשב לקביעת הפלוטיפים מותאמת לעדר הבקר.

Baruch et al., Efficient inference of haplotypes from genotypes on a large animal pedigree. *Genetics*. 2006 172:1757-65.

כן פורסם מאמר המתאר תוכנת מחשב לקביעת הפלוטיפים בריצוף ישיר.

Seroussi & Seroussi, TraceHaplotyper: using direct sequencing to determine the phase of an indel followed by biallelic SNPs. *Biotechniques*. 2007 43:452-56.

**מסקנות והמלצות:** קביעת הפלוטיפים היתה פתח לשלב מחקרי נוסף של מיפוי הגן האחראי על בסיס חוסר שיווי משקל באוכלוסיה. נבחנו האלגוריתמים הקיימים לצורך האמור וניסינו להתאימם לשימוש בעדר הבקר. התוכנה WHAP\PLINK נמצאה כאלטרנטיבה המועדפת. התוכנות שנכתבו ונבחרו נבחנו ויושמו בנתוני אמת, במאמצים מחקריים מקבילים לאיתור גנים כמותיים לתכונות חלב בכרומוזומים 6 ו-7 של בקר. תרומתם התבטאה, בין השאר, במאמר שהתקבל ופרסום ומסכם את חשיבות האנליזה בעזרת הפלוטיפים לזיהוי הגן הכמותי בכרומוזום 6.

Seroussi, Concordance test emerges as a powerful tool for identifying quantitative trait nucleotides: lessons from BTA6 milk yield QTL. *Animal Genet*. 2008 in press.

### **2.1.2 Abstract**

**The problem:** During the bovine genome scan we mapped loci for quantitative economical traits of the Israeli dairy herd. Identification of the underlying genes and their responsible genetic variation are a complex and expensive missions. Using haplotype based selection, it is possible to select for the quantitative gene even when the actual underlying variation has not been identified. Haplotype is a set of single nucleotide polymorphisms (SNPs) on a single chromatid that are statistically associated. Such genetic markers are wide spread and are amenable to automation. Haplotypes of genetic markers may be reconstructed using pedigree data. Although the reconstruction is based on inference of clear and sequential rules it is too complex for human calculation when dealing with numerous markers on the population level.

**Aims:** To implement an algorithm for haplotype reconstruction in software useful for the dairy herd and to examine algorithms for population-wide linkage-disequilibrium (LD) mapping that uses the reconstructed haplotypes.

**Methods and implementation:** DNA was collected from 716 bulls with genetic evaluations based on their daughter performance. This sire population has complex pedigree relations that allow inference of the common haplotypes in the Israeli herd. SNPs were genotyped in critical regions on BTA6 and BTA7. Computer programs were written in C++ and PERL.

**Results:** A computer program for large scale haplotype reconstruction was developed and published: Baruch et al., Efficient inference of haplotypes from genotypes on a large animal pedigree. *Genetics*. 2006 172:1757-65. A computer program for the determination of haplotype using direct sequencing was written and published, as well: Seroussi & Seroussi, TraceHaplotyper: using direct sequencing to determine the phase of an indel followed by biallelic SNPs. *Biotechniques*. 2007 43:452-56.

**Conclusions:** Reconstruction of haplotypes has been the basis for the LD mapping that followed. We tested algorithms for studying population wide LD for identification of quantitative genes and adjusted them for the use in the dairy herd. The program WHAP/PLINK was found as the best alternative. We implemented the algorithms that were written and selected in the analysis of real data obtained from parallel efforts aiming at the identification of quantitative genes on BTA6 and BTA7. Their contribution was exemplified in the publication that point to the importance of haplotype analysis to the identification of the quantitative gene on BTA6: Seroussi, Concordance test emerges as a powerful tool for identifying quantitative trait nucleotides: lessons from BTA6 milk yield QTL. *Animal Genet*. 2008 *in press*.

## 2.2 מבוא

במהלך סריקת הגנום מופו אתרים לתכונות כלכליות כמותיות (QTL) בבקר לחלב בישראל. לאתרים אלה חשיבות רבה בטיפול העדר. זיהוי מדויק של הגנים האחראים ואיתור השונות הגנטית בהם הם משימות מורכבות ויקרות. ניתן לבצע סלקציה לגן כמותי בהעדר זיהוי של הגן עצמו באמצעות סלקציה להפלוטיפ, הכולל את האלל הרצוי. הפלוטיפ הוא רצף מסוים של אללים על פני הכרומוזום המאפיין קבוצת סמנים גנטיים קרובים התוחמים את ה-QTL. למטרה זאת משמשים סמנים מסוג של שינוי בבסיס יחיד (SNP) הנפוצים בגנום, הנוחים לזיהוי ולקביעת גנוטיפים ממוכנת. אנליזה במבנה משפחתי מאפשרת קביעת הפלוטיפים באוכלוסיה הנבדקת. נאסף חומר גנטי מ-716 פרים עם אומדני הורשה שנקבעו על בסיס מבחן צאצאים. אוכלוסיית פרים זאת היא בעלת קשרי משפחה מסועפים המאפשרים מעקב אחר ההפלוטיפים השכיחים בעדר הישראלי.

למרות ששיטות ההסקה הן סדרתיות ומובנות, הן מורכבות למדיי לביצוע אנושי בפרק זמן סביר עבור מספר סמנים רב באוכלוסייה. לכן, לפיתוח תוכנת מחשב לקביעת הפלוטיפים של כל פרט חשיבות מכרעת ומורכבות גבוהה.

על תוכנה לענות לדרישות הבאות:

1. פענוח מרבי של ההפלוטיפים במדגם של מאות פרטים, בעלי קשרי משפחה מסועפים. 2. השלמת מידע גנטי חסר של אמהות הפרים ושל גנוטיפים חסרים בפרים. 3. דווח על סתירות, במידה שקיימת אי התאמה בין גנוטיפים לפי חוקי ההורשה של מנדל, כדי לאפשר למשתמש לאתר את מקור השגיאה ולקבוע דרך טיפול מתאימה.

בבסיס התוכנה יישם האלגוריתם שעונה חלקית לדרישות ושהוצע ע"י Li J., and Jiang T. (2003). התוכנה תפותח בהנחה של חוסר שחלוף באזור הקריטי. התוכנה תבדק ע"י הדמיה של מבנה נתונים עם הפלוטיפים קבועים מראש ונתוני אמת של סמנים באוכלוסיית הבקר הישראלי.

בהמשך לפיתוח התוכנה, ייבחן שימוש בהפלוטיפים שנקבעו לצורך צמצום האזור הקריטי של גן כמותי עד לאזור התוחם גנים בודדים ע"י מיפוי בעזרת אי-שיווי משקל בתאחיזה (Linkage disequilibrium -LD) (mapping). יבחנו מספר אלגוריתמים קיימים לצורך זה ותאופיין האנליזה הסטטיסטית האופטימלית, במבנה הנתונים הקיים בידנו.

## 2.3 מטרת המחקר

- לפתח תוכנת מחשב לקביעת הפלוטיפים המותאמת לעדר הבקר.
- בחירה או התאמה של אלגוריתם למיפוי בעזרת הפלוטיפים של הגן האחראי על בסיס חוסר שיווי משקל באוכלוסייה.

## 2.4 חשיבות המחקר

השגת המטרה עיקרית של תוכנית זאת שהיא לפתח תוכנת מחשב מותאמת לעדר הבקר, תאפשר סלקציה לתכונה הכמותית גם בהיעדר זיהוי של הגנים האחראים וניצול המידע שהצטבר בתוכניות המחקר של מפוי עדין של אתרים אלה (כרומוזומים 6, 7, ו-2, 364-0175, 364-0177, 364-0181, בהתאמה). השגת המטרה הנוספת של התוכנית שהיא בחירה של אלגוריתם למיפוי הגן האחראי על בסיס חוסר שיווי משקל באוכלוסייה תייעל את איתור הגן האחראי והמעקב אחריו.

## 2.5 פירוט עיקרי הניסויים

נכתבה תוכנית מחשב בשפת התכנות ++C המאפשרת שיחזור הפלוטיפים בקנ"מ נרחב באוכלוסיית הפרים להזרעה בישראל. פורסם המאמר המתאר את האלגוריתם שבבסיס התוכנה:

Baruch E., Weller J.I., Cohen M., Ron M., Seroussi E. (2006) Efficient Inference of Haplotypes from Genotypes on a Large Animal Pedigree. *Genetics* **172**, 1757-1765.

נכתבה תוכנית מחשב בשפת התכנות Perl המאפשרת שיחזור הפלוטיפים בעזרת רצוף ישיר. ופורסם מאמר המתאר את האלגוריתם שבבסיס התוכנה.

Seroussi, Y., Seroussi, E. (2007) TraceHaplotyper: Using direct sequencing to determine the phase of an indel followed by biallelic SNPs. *BioTechniques* **43**, 452-456.

נבחנה התאמה של מספר תוכנות לאיתור גנים כמותיים, בעזרת אנליזה סטטיסטית של הפלוטיפים, על בסיס חוסר שיווי משקל באוכלוסייה.

תוכנות שתיאורן העיד על התאמה אפשרית לשימוש בעדר הבקר הותקנו והשוותה יכולתן במבחני אסוציאציה. בהשוואה נעשה שימוש בנתוני גנוטיפים של מיפוי גנים כמותיים בכרומוזום 6 ו-7 של בקר.

התוכנות שנבחנו:

שם	מובאה	מאפיינים עקריים
TreeLD	Zollner et al. 2005	פותחה עבור נתונים אנושיים. מבוססת על הרעיון שעדות לזיקה לאתר בגנום נמצאת בעץ התורשתי של הפלוטיפים שנדגמו.
Tassel	Yu et al., 2006	פותחה עבור נתונים של תירס. אנליזה של זיקה לתכונות יבול, דפוסי אבולוציה ו-LD, תוך שימוש במודל מעורב לאנליזה של אסוציאציה גנטית, בהתחשבות ברמות רבות של קרבה.
WHAP	Purcell et al., 2007 <a href="http://pngu.mgh.harvard.edu/~purcell/">http://pngu.mgh.harvard.edu/~purcell/</a>	פותחה עבור נתונים אנושיים. מבצעת אנליזה מבוססת הפלוטיפים של תכונות כמותיות עם סמני SNP ו/או סמנים מיקרוסטיליטיים באוכלוסייה או במבנה משפחתי. מציעה טווח רחב של מבחנים סטטיסטיים. לאחרונה הופסקה התמיכה בתוכנה זאת בכוונה לשלבה בתוכנה *PLINK.

\* PLINK נועדה בעיקרה לניתוח בקנ"מ גנומי נרחב של סמנים (למשל נתונים מ-BeadChip), היא משתלבת עם התוכנה הפופולרית HAPLOVIEW (<http://www.broad.mit.edu/haploview/haploview>) לתצוגה אופטימלית של התוצאות ומותאמת למערך הכרומוזומאלי של בקר.

התוכנה WHAP נמצאה כאלטרנטיבה הנוחה ביותר לשימוש, במקרה שקיים חיזוי מדויק של הפלוטיפים, כמודגם בפסקה הבאה.

### 2.5.1 ישום הפלוטיפים לבחינת זיקה גנטית בין נתוני הפרים והגנים כמותיים בכרומוזום 7 של בקר.

עבור אוכלוסייה של 716 פרי הזרעה ישראליים, נקבעו גנוטיפים לעשרים סמני SNP, באזור בין כ-50 אלף בסיסים בכרומוזום 7. נתונים אלה, בצרוף המבנה המשפחתי, הוזנו לתוכנת שיחזור הפלוטיפים שפתחנו (LSPH). הפלט העיד שהאללים של הסמנים האמורים יוצרים כ-20 הפלוטיפים, מתוך כמיליון הצרופים האפשריים.

בשלב שני נופו 73 השגיאות שנבעו מחוסר התאמה בין הורים לצאצאיהם ושהתגלו. ניתן להעריך שרמת הדיוק בקביעת הגנוטיפים היא כ-95%, בין השאר, ע"פ סמן שנבחן פעמיים בעזרת פריימרים שונים (39 אי התאמות ב-300 חזרות).

בשלב שלישי שובצו והושלמו האללים הבלתי פתורים לפי העמדות התואמות את ההפלוטיפים שקיומם הוכח, ונכונות הפתרון נבחנה בעזרת LSPH. פתרון זה משער מבנה הפלוטיפים מושלם עבור 20 הסמנים ל-676 פרים עם רשומות יבול.

נתוני היבול והפלוטיפים הוזנו עבור פרי הזרעה אלה, כ"א כמשפחה נפרדת, לתוכנת WHAP כמתואר בפלט שבטבלה 1. יש לציין שזהות הפרים במדגם הפרים בו שוחזרו הפלוטיפים מלאים אינה חופפת במלואה לזהות הפרים במדגם לו נקבעו גנוטיפים. סה"כ 9 הפלוטיפים המייצגים 96% מהמדגם השתתפו בחישוב.

**טבלה 1.** שורת הפקודה (אפור) ופלט לדוגמה מתוכנת WHAP : אנליזה של סבירות (צהוב) לזיקה מקרית בין אחוז החלבון בחלב בנות פרים להפלוטיפים של 20 סמני SNP בכרומוזום 7 של בקר. הקלט כלל שלושה קבצים: קובץ בין 676 שורות (BTA7\_%protein.ped), בכל שורה נתוני זהוי של הפר, המדד הגנטי של % חלבון בחלב בנותיו, וגנוטיפים של 20 סמנים בסדר תואם להפלוטיפ המשוער, כאשר הנוקלאוטידים A, C, G, T מיוצגים ע"י הספרות 1, 2, 3, ו-4, בהתאמה; קובץ מפה הכולל רשימה של 20 הסמנים ומיקומם (BTA7\_%protein.map) וקובץ נתונים המציין את סוג התכונה הנבדקת ושמות הסמנים (BTA7\_%protein.dat). מומלץ להפעיל את האופציה לחזרה על החישוב (--repeat 100) משום שבאלגוריתם התוכנה מתבצע חיפוש רנדומלי אחר הפתרון שעלול לא להתכנס בהרצה יחידה.

```
C:\WHAP>whap --file BTA7_%PROTEIN --repeat 100

WHAP! | v2.09 | 26/Oct/05 | S. Purcell, P. Sham |
spurcell@pngu.mgh.harvard.edu
676 individuals w/out parents. 0 individuals with parents. Continuous trait:

624 of 676 individuals/trios are informative
Hap          Freq      Alt(B)      Alt(W)      Null(B)  Null(W)
---          -
11412433323311222432  0.438      0.000      0.000 [1]      0.000      0.000 [1]
24212213323231222432  0.161     -0.013     -0.013 [2]      0.000      0.000 [1]
11412233323211242432  0.128      0.005      0.005 [3]      0.000      0.000 [1]
24244213323231222432  0.113      0.061      0.061 [4]      0.000      0.000 [1]
11412433133211222432  0.058      0.009      0.009 [5]      0.000      0.000 [1]
11412433323213424244  0.048      0.004      0.004 [6]      0.000      0.000 [1]
24212233323211222432  0.022     -0.018     -0.018 [7]      0.000      0.000 [1]
11412433323211242432  0.019     -0.011     -0.011 [8]      0.000      0.000 [1]
24212231321231222432  0.015     -0.011     -0.011 [9]      0.000      0.000 [1]
---          -
                                -1531.713      -1438.023

Proportion of haplotypes covered = 0.959
LRT = 93.690
df = 8
p = 8.27e-17
```

**טבלה 2.** האנליזה ב-WHAP של השפעות של הפלוטיפים בכרומוזום 7 על נתוני יבול.

בטורים שמייצגים תכונות מובהקות הודגשו הקיצוניות הרצויות (כחול) ושאינן רצויות (אדום) לחקלאי.

ערכי ההשפעה המקסימליים לתכונות היבול: כמות ו-% השומן, כמות ו-% החלבון ואינדקס PD07 וכן ההפלוטיפ הנלווה אליהן מודגשים בצהוב.

ערכי ההשפעה המקסימליים לכמות חלב וההפלוטיפ הנלווה אליה מודגשים בתכלת.

ערך ההשפעה המקסימלי לפוריות בנות וההפלוטיפ הנלווה אליה מודגשים בוורוד.

#	Haplotype	Freq	Effects relative to the common haplotype									
			Milk	Fat	% fat	Protein	% protein	PD07	Fertility	Twining	Persistency	SCS
1	11412433323311222432	0.438	0.0	0.000	0.000	0.000	0.000	0.0	0.000	0.000	0.000	0.000
2	24212213323231222432	0.161	-28.4	-3.605	<b>-0.024</b>	-2.311	-0.013	-113.6	-0.388	0.136	-0.114	<b>0.038</b>
3	11412233323211242432	0.128	45.3	3.258	0.016	1.867	0.005	41.5	-0.637	0.243	-0.021	0.006
4	24244213323231222432	0.113	-65.5	<b>5.202</b>	<b>0.070</b>	<b>4.508</b>	<b>0.061</b>	<b>122.3</b>	<b>-1.062</b>	0.038	-0.455	-0.036
5	11412433133211222432	0.058	17.1	-0.755	-0.013	1.480	0.009	60.8	0.336	0.285	0.091	<b>-0.048</b>
6	11412433323213424244	0.048	<b>90.2</b>	0.733	-0.023	3.372	0.004	111.2	-0.087	<b>0.804</b>	-0.043	-0.047
7	24212233323211222432	0.022	<b>-187.3</b>	<b>-6.554</b>	0.000	<b>-7.744</b>	<b>-0.018</b>	<b>-214.1</b>	<b>0.831</b>	<b>-0.802</b>	-0.495	-0.037
8	11412433323211242432	0.019	-129.8	-4.741	-0.001	-5.288	-0.011	-211.6	-0.492	-0.265	-0.586	-0.001
9	24212231321231222432	0.015	38.5	2.551	0.010	0.007	-0.011	37.6	0.531	0.502	0.727	0.023
	Probability	0.959*	$9.0 \times 10^{-04}$	$6.3 \times 10^{-10}$	$7.6 \times 10^{-09}$	$1.0 \times 10^{-11}$	$8.3 \times 10^{-17}$	$2.4 \times 10^{-09}$	$8.2 \times 10^{-07}$	$4.0 \times 10^{-04}$	0.34	0.04

\* Proportion of haplotypes covered

האנליזה ב-WHAP של השפעות של הפלוטיפים על נתוני יבול (טבלה 2). מעידה על חשיבות חקלאית מובהקת של אזור כרומוזומלי זה לכל תכונות היבול להוציא ספירת התאים הסומטיים. אותו שלושה הפלוטיפים ראויים לציון: הפלוטיפ מס. 4 (צהוב, טבלה 2) משויך לעליה בכמות ובריכוז השומן והחלבון בחלב והשפעה חיובית ניכרת על אינדקס הטיפוח (PD07). להשפעות חיוביות אלה נלווה אפקט שלילי ניכר על פוריות. הפלוטיפ האמור נדיר יחסית באוכלוסייה ועל כן יש מקום לשקול השבחה מכוונת סמנים כדי להגביר את שכיחותו בעדר החלב. הפלוטיפ מספר 6 (תכלת, טבלה 2) משויך לערכים מקסימליים של השפעות על תפוקת החלב ושכיחות לידות תאומים. קשר חיובי בין שתי התכונות הוצע בעבר על ידי קבוצת מחקר שבחנה את תכונות אלה בכרומוזום 7 (Cruickshank et al., 2004). לידת תאומים היא תכונה בלתי רצויה לחקלאי עקב הסיבוכים הרפואיים הנלווים ואיילונות. אי לכך ספק אם יש להעדיף בהשבחה את נושאי הפלוטיפ 6 האמור. הפלוטיפ מספר 7 הוא בעל ההשפעה החיובית ביותר על פוריות, אולם להשפעה זאת נלוות פגיעה ברוב תכונות היבול המסתכמת בהשפעה שלילית קיצונית על אינדקס הטיפוח. יש לציין כי בניגוד להפלוטיפים הקודמים הפלוטיפ זה לא ניתן לאיתור על ידי סמן מנחה יחיד. דבר זה מעיד על חשיבות אנליזה של מספר סמנים קשורים (הפלוטיפ) וכי אנליזה לפי הפלוטיפים צפויה להיות עמידה יותר בפני טטיות מקריות (Akey et al., 2001).

#### 2.5.2 ישום הפלוטיפים לבחינת זיקה גנטית בין נתוני הפרים והגנים כמותיים בכרומוזום 6 של בקר.

התוכנות שנכתבו ובחרו נבחנו ויושמו במחקר לאיתור גנים כמותי עקרי לתכונות חלב בכרומוזומים 6 של בקר (Baruch et al., 2006). תרומתם התבטאה, בין השאר, במאמר שהתקבל ופרסום ומסכם את חשיבות האנליזה בעזרת הפלוטיפים לזיהוי הגן הכמותי בכרומוזום 6, ומצורף כנספח לדו"ח זה. Seroussi, Concordance test emerges as a powerful tool for identifying quantitative trait nucleotides: lessons from BTA6 milk yield QTL. *Animal Genet.* 2008 *in press.*



Akey, J., L. Jin and M. Xiong. 2001. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J Hum Genet* **9**:291-300.

Baruch E., Weller J.I., Cohen M., Ron M., Seroussi E. (2006) Efficient Inference of Haplotypes from Genotypes on a Large Animal Pedigree. *Genetics* **172**, 1757-1765.

Cohen-Zinder M, Seroussi E, Larkin DM, Looor JJ, Everts-van der Wind A, Lee JH, Drackley JK, Band MR, Hernandez AG, Shani M, Lewin HA, Weller JI, Ron M. (2005) Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* **15**:936-944 .

Cruickshank J, Dentine MR, Berger PJ, Kirkpatrick BW. (2004) Evidence for quantitative trait loci affecting twinning rate in North American Holstein cattle. *Anim Genet.* **35**:206-212.

Li J., and Jiang T. (2003) Efficient Inference of Haplotypes from Genotypes on a Pedigree, *Journal of Bioinformatics and Computational Biology*, **1**: 41-69.

Purcell S, Daly MJ, Sham PC. (2007) WHAP: haplotype-based association analysis. *Bioinformatics*, **23**:255-256

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, **81**.

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES. (2006) A Unified Mixed-Model Method for Association Mapping Accounting for Multiple Levels of Relatedness. *Nature Genetics*, **38**:203-208.

Zollner S, Wen X, Pritchard JK. (2005) Association mapping and fine mapping with TreeLD. *Bioinformatics*, **15**:3168-3170.

### 3. סיכום עם שאלות מנחות

נא לענות על כל השאלות, בקצרה ולעניין, ב 3 עד 4 שורות מכסימום לכל שאלה (לא תובא בחשבון חריגה מגבולות המסגרת המודפסת).

שיתוף הפעולה שלך יסייע לתהליך ההערכה של תוצאות המחקר. הערה: נא לציין הפנייה לדו"ח אם נכללו בו נקודות נוספות לאלה שבסיכום.

<p><b>מטרות המחקר לתקופת הדו"ח תוך התייחסות לתוכנית העבודה.</b></p> <p>1- לפתח תוכנת מחשב לקביעת הפלוטיפים המותאמת לעדר הבקר. 2- בחירה של אלגוריתם למיפוי בעזרת הפלוטיפים של הגן האחראי על בסיס חוסר שיווי משקל באוכלוסיה.</p>
<p><b>עיקרי הניסויים והתוצאות שהושגו בתקופה אליה מתייחס הדו"ח.</b></p> <p>נכתבה תוכנית מחשב בשפת התכנות Perl המאפשרת שיחזור הפלוטיפים בריצוף ישיר. נבחרה התוכנה WHAP/ PLINK מבין האלטרנטיבות לאנליזה בעזרת הפלוטיפים של הגן האחראי על בסיס חוסר שיווי משקל באוכלוסיה. התוכנות שנכתבו ונבחרו נבחנו ויישמו בנתוני אמת, במאמצים מחקריים מקבילים לאיתור גנים כמותיים לתכונות חלב בכרומוזומים 6 ו-7 של בקר.</p>
<p><b>המסקנות המדעיות וההשלכות לגבי יישום המחקר והמשכו. האם הושגו מטרות המחקר בתקופת הדו"ח.</b></p> <p>מטרות המחקר (1, 2) לתקופת הדו"ח הושגו.</p>
<p><b>הבעיות שנתרו לפתרון ו/או השינויים שחלו במהלך העבודה (טכנולוגיים, שיווקיים ואחרים); התייחסות המשך המחקר לגביהן, האם יושגו מטרות המחקר בתקופה שנתרה לביצוע תוכנית המחקר.</b></p> <p>תוכנית המחקר הסתיימה.</p>
<p>האם הוחל כבר בהפצת הידע שנוצר בתקופת הדו"ח - <b>יש לפרט</b>: פרסומים – כמקובל בביבליוגרפיה, פטנטים - יש לציין מס' פטנט, הרצאות וימי עיון - יש לפרט מקום ותאריך.</p>
<p>Baruch, E., Weller, J.I., Cohen, M., Ron, M., Seroussi, E. (2006) Efficient Inference of Haplotypes from Genotypes on a Large Animal Pedigree. <i>Genetics</i> <b>172</b>, 1757-1765.</p> <p>Seroussi, Y., Seroussi, E. (2007) <i>TraceHaplotyper</i>: Using direct sequencing to determine the phase of an indel followed by biallelic SNPs. <i>BioTechniques</i> <b>43</b>, 452-456.</p> <p>Seroussi, E. (2008) Concordance test emerges as a powerful tool for identifying quantitative trait nucleotides: lessons from BTA6 milk yield QTL. <i>Animal Genet. in press</i>.</p>
<p>פרסום הדו"ח: אני ממליץ לפרסם את הדו"ח: (סמן אחת מהאופציות)</p> <p>ללא הגבלה (בספריות ובאינטרנט) ←</p>