

יישום הערכות גנומית לעגלות ולפרות חולבות

Application of genomic evaluations for heifers and dairy cows

מורן גרשוני ואיל סרוסי, מכון לחקר בע"ח, המנהל למחקר חקלאי – מכון וולקני.

תקציר.

שיטות יעילות לברירה גנומית של פרי הזרעה, על בסיס קביעת נרחבת של גנוטיפים עם טכנולוגית השבב הגנטי (Illumina BovineSNP50TM וניגזרותיו), מיושמות כיום בהיקף נרחב במשק הבקר לחלב העולמי ובישראל. עם זאת, יישום אפקטיבי דורש סדרת אימון גדולה, שאינה זמינה עבור העדר הישראלי. בעת ביצוע תחזיות גנומיות באמצעות רגרסיה של פנוטיפים לשבבים גנטיים עם מערכי SNPs בצפיפות גבוהה, מספר הסמנים (p) עולה בהרבה על מספר הדגימות הזמינות (n). רגרסיה נרחבת כזאת, כאשר $n < p$ הופכת לבעיה סטטיסטית, כך שרגרסיה זאת דורשת שיטות לברירה או לייצוג מופחת של מספר הסמנים. מצד שני, נצברה בנתוני ספר העדר, כמות משמעותית ($> 9,000$) של רשומות השבב הגנטי עבור עגלות ופרות חולבות, כולל עתודות ופרות רגילות של משקים שאימצו טכנולוגיה זאת, בעיקר כדי לאמת הורות וגם לקבלת דירוג להוצאה. השערת המחקר היא ששילוב נתונים אלה עם נתוני פרי ההזרעה ישפר את יכולת החיזוי על בסיס נתונים גנומיים ויאפשר ניבוי ערכי טיפוח (gEBVs) מהימנים יותר הן לפרים והן לפרות. על בסיס שילוב זה, מטרות המחקר היו: (1) לגבש אלגוריתם לקבלת הערכות גנומיות משופרות לתוכנית ההשבחה של בקר לחלב, אשר בנוסף תשמשנה בידי הרפתן ככלי מיעץ בהסדרת סדר היציאה של פרות מהרפת ובקביעת הכדאיות בהזרעתן; (2) להעריך את הכדאיות הכלכלית ביישום כאמצעי להשבחה ולשיפור הממשק שנגזר מאימות הורות; (3) להכשיר כוח אדם מתאים ליישום התחום ומחקרו.

בשלב הראשון גויס מתמחה בת"ר דוקטור, ד"ר אריה קורזון שנקלט כיום במערך ספר העדר, והוביל את המחקר תוך יישום תוכנת מחשב חדשנית (APEX). תוכנה זו מסוגלת לחשב ערכי טיפוח גנומיים בשיטה חד-שלבית (ssGBLUP) (gEBVs), דבר המאפשר שילוב פרות ופרים יחד. במתווה זה הדגמנו כי שילוב גנוטיפים של פרות הינו בעל ערך לבניית מערכי אימון עם כוח סטטיסטי משמעותי, המאפשרים חישוב של ערכי טיפוח גנומיים אמינים לתוכנית ההשבחה בישראל. מצאנו כי נתונים ישנים הכוללים גנוטיפים מועטים אינם משפרים את החיזוי. לעומת זאת, הגבלת השימוש בסמנים לגנוטיפים עדכניים שיפרה את התחזיות של ערכי הטיפוח הגנומיים. בנוסף, כווננו הפרמטר של המשקל הפוליגמי וטווח הזמן של דגימת הנתונים השפיע על דיוק של קביעת של ערכי הטיפוח הגנומיים ומיתון הערכת יתר שלהם. בעבודה משלימה שמטרתה היתה בחינת אסטרטגיות לשיפור נוסף בהערכות היתר שנמצאו במודל ssGBLUP עבור תנבות שומן וחלבון, באמצעות שילוב של ערכי טיפוח מנוקי-רגרסיה (Deregressed Proofs; DRP) ומידע בינלאומי שהופק ב- Interbull. בעבודה זו נמצא כי שילוב DRP ונתוני האינטרבול וכן שילוב משקלי תרומת הסמנים הביאו לשיפור מובהק בדיוק התחזית תוך שיפור, עד ביטול, הערכות היתר. לבסוף, נבחנו תחזית ריבוי בשארות על בסיס הנתונים הגנומיים מפרות ופרים. נבדקו מספר אסטרטגיות, ונמצא כי שימוש במטריצת הייחוסים הגנומית (GRM) מפיקה מחד תחזיות גבוהות באופן מובהק מהשיטה הנהוגה היום המבוססת על תיעוד שושלות (קורלציה של 0.7 ל- GRM לעומת 0.56 עבור תיעוד שושלות), ובנוסף, הראנו כי יישום חישוב ריבוי בשארות על בסיס GRM הינו פשוט יחסית עם עלויות חישוביות מתונות. תוצאות אלו מדגימות כי הפעלת הערכות גנומיות תוך שילוב גנוטיפים מפרות, אשר מספרם עולה בכל שנה, יכולה לייצר אומדנים מדויקים בדומה להערכות שמבוצעות הנשענות על נתוני העדר ההולנדי (CRV). תוצרי המחקר יובילו להאצת ההתקדמות הגנטית במשק הבקר לחלב, שיפור של הממשק לניהול פרות הרפת ושל התחרותיות של ייצור החלב בישראל לעומת חלופות בחו"ל, וכן סייעה בהכשרת הדור הבא בענף.

1. מבוא ותיאור הבעיה:

ערכי טיפוח גנומיים (gEBV) משמשים כדי לחזות את ערכם של פרטים בעלי גנוטיפים לצורך תכנית ההשבחה [Meuwissen 2001]. לצורך זה, קביעה נרחבת של גנוטיפים עם טכנולוגית השבב הגנטי (Illumina BovineSNP50TM וניגזרותיו), מיושמת כיום בהיקף נרחב במשק הבקר לחלב העולמי ובישראל. בדרך כלל, האלגוריתמים של קביעת ערכי טיפוח גנומיים אינם מתייחסים לכך שאוכלוסיית הייחוס, המשמשת להערכת השפעות הסמנים על הגנוטיפים, מורכבת ממדגם מוטה של פרי הזרעה נבחרים, מה שמוביל אולי להטיה כמו גם לירידה ברמת הדיוק. הוצעה נוסחה לחישוב מקורב של המהימנות ושל ההטיה של ערכי הטיפוח הגנומיים, שמביאות בחשבון את הברירה המוטה של פרי נבחנים עם גנוטיפ מבין כלל הפרים [Mäntysaari 2010]. בשיטה זאת, ההבדלים בין הממוצעים וסטיות התקן של ערכי הטיפוח של כלל הפרים משמשים להערכת השיעור של ההטיה. לאחר מכן, הבדל בברירה ועוצמתה מחושבים לפי המקובל לגנטיקה כמותית [Falconer 1996], כדי לחשב את המהימנות של ערכי הטיפוח הגנומיים ואת הטייתם האמורות [Mäntysaari 2010]. עם זאת, השונות הגנטית האמיתית מופחתת לא רק על ידי עוצמת הברירה אלא גם על ידי מהימנותם של ערכי הטיפוח המחושבים בקירוב [Falconer 1996]. מהימנות זאת משתנה לפי התכונה ובין פרים לפרות. באוכלוסיות בקר לחלב, עוצמת הברירה גבוהה יותר בזכרים מאשר בנקבות [Schaeffer 2006], מה שמצביע על כך שניתן להוסיף לאוכלוסיית הייחוס פרות עם פחות הטיה ולהקטין בכך את אובדן הדיוק.

בעבר ניתן לפרים ערך טיפוחי (EBV) ללא מידע גנומי ורק על בסיס קשרים משפחתיים ופנוטיפ (ערך טיפוחי מסורתי). היישום כלל מודל ליניארי מעורב הנקרא BLUP שפותח ע"י Henderson בשנות ה-50 של המאה הקודמת (Freeman et al. 1991), ושילב מתן אומדנים לאפקטים אקראיים. בהמשך פותחו שתי שיטות עיקריות למתן אומדן עבור ערכי הטיפוח בעזרת נתונים גנומיים (GEBV). VanRaden וחבריו (2008) הציעו שיטה הנקראת genomic BLUP (gBLUP) המבוססת על יצירת מטריצת יחסים בין הפרטים על בסיס המידע הגנומי. מטריצה זו המיוצגת על ידי האות G באה במקום מטריצת היחסים בין הפרטים המיוצגת ע"י האות A שנקבעת על בסיס רישום קשר משפחתי בלבד. שיטה זו מכונה השיטה הרב שלבית היות ויש בה מספר שלבים על מנת שניתן יהיה לשלב את המידע הגנומי באנליזה. בשלב הראשון מחשבים את הEBV המסורתי עבור כל השושלת. בשלב השני ביצועי הפרות משמשים על מנת לתת ערך טיפוחי לאבות על ידי אחת משתי שיטות daughter yield deviations (DGD) או deregressed evaluations (DD). בשלב השלישי הקשר בין המידע הגנומי (G) נבחן אל מול DD או DGD ועבור כל פרט בעל נתונים גנומיים ניתן ערך טיפוחי direct genomic values (DGV). ניתן לשלב עם הDGV עם מידע של ממוצעי ההורים או עם EBV מסורתי המחושב על בסיס קשר משפחתי וכך מתקבל GEBV (Lourengo et al. 2020). Misztal וחבריו (2009) הציעו את השיטה החד שלבית. שיטה זו משלבת פנוטיפים, נתוני שושלת וגנוטיפים באנליזה אחת לצורך אומדן ערכים טיפוחיים גנומיים (ssGBLUP). במקום מטריצת היחסים G המשמשת באנליזה הרב שלבית הכוללת רק מידע גנומי, משתמשים במטריצה היברידית (H) שכוללת בתוכה גם מידע גנומי וגם מידע שושלתי.

שילוב פרות באנליזה רב שלבית יחסית מסובך (Koivula et al. 2016) ביחס לשיטה החד שלבית למרות שבמשך השנים נעשו התאמות לשילוב פרות בשיטה זו. Wiggans וחבריו (2011) הציעו תיקון בערכי הטיפוח של הפרות היות והם היו מנופחים ביחס לערכי הפרים, לאחר תיקון זה נטען ששילוב פרות באנליזה הוא רצוי. עבודה שנעשתה לאחרונה (De Oliveira et al. 2019) הציעה שיטה אפשרית לשילוב פרות באנליזה הרב שלבית שכוללת הסרה של מידע על

דוח מסכם 362-0774 (יוני 2026)

פרות עם נתונים גנומים וממוצע הורים בחישוב של GEBV עבור הפרים וזאת כדי להימנע מבעית של דגימה כפולה של אותה פרה בנתונים.

שילוב פרות באנליזה החד שלבית אינו דורש התאמות מיוחדות ונחשב לפשוט יחסית (Koivula et al. 2016). מספר תוכנות משמשות לאנליזה חד שלבית. התוכנה BLUPF90 שפותחה ע"י צוותו של Misztal שופרה למתן ביצועים גבוהים ונכנסים בה שיפורים חדשים כמה פעמים בשנה (Lourenco et al. 2020). הגרסה החינמית של BLUPF90 מאפשרת אנליזה של 25,000 חיות עם גנוטיפ (Lourenco et al. 2020). התוכנות MiXBLUP ו MiX99 גם יכולת לשמש לצורך דומה (Alkhoder et al. 2022) התוכנה MiXBLUP משתמשת בנגזרת מסוימת של השיטה החד שלבית (ssSNPBLUP) ונעשה בה שימוש מוצלח לאחרונה באנליזה שכללה פרות בגרמניה (Alkhoder et al. 2019). בעבר נעשה ניסון להכניס פרות לאנליזה בישראל (Lourenco et al. 2014), אך מאז גדל מסד הנתונים של הפרות שעברו ריצוף למספרים גבוהים בהרבה מה שמצדיק בדיקה נוספת.

קביעת גנוטיפים לפרות הופכת נפוצה יותר ככל שעלותה יורדת. יתרון מובהק של שימוש בנתונים אלה הוא תיקון רשומות ההורות ברפת, שחיוניות לניהול תכנית ההזרעות תוך צמצום של הריבוי בשארות (inbreeding). לזיווג קרובים יש השפעה שלילית שנובעת הן מגנים פגומים רצסיביים שבאים לכדי ביטוי רק במצב הומוזיגוטי והן מירידה בתפקוד של מערכת החיסון, בה המגוון של האללים חיוני ליכולת תגובה נגד מגוון רחב של פתוגנים. בארה"ב הנזק הוערך בכ- \$25 הפחתה במדד התפוקה (Net Merit), לכל תוספת של אחוז גידול בשארות [VanRaden 2017]. בישראל, על בסיס פרופיל גנטי, חושב שיעור של כ- 5% טעות ברשומות ההורות של פרות, שבעיקר יחסו לטעות בזיהוי הפר המזריע [Ron 1996]. חוקרים אלה גם העריכו שבטווח הארוך (20 שנה), שיעור שגיאות זה מוביל לנזק שווה ערך למיליוני ₪ (2.4 מיליון \$) לתוכנית ההשבחה בישראל. תוצאות אלה הביאו להקפדת יתר ברישומי המזריעים, אבל לפי הרישומים והמעקב ברפת ולקני, על בסיס נתוני השבב הגנטי בפרות, נותר שיעור הרשומות השגויות דומה. זאת מסיבות מגוונות, כולל קשיים בשיוך האם במקרה של המלטות סמוכות, וכן טעויות אנוש ברישום (תוצאות לא מוצגות). בניגוד לשימוש בפרים, בחישוב ערכי טיפוח גנומיים בפרות, אותם פרטים מקבלים גנוטיפ וגם פנוטיפ (חישוב ישיר - Single step), במקום גנוטיפ של ההורה ופנוטיפ של צאצאים. לכן הוספת גנוטיפים של נקבות ורשומותיהן הפנוטיפיות לאוכלוסיית הייחוס הקיימת לאבות צפויה להגביר את מהימנותן של ערכי הטיפוח הגנומיים, ועלייה במהימנות תוביל לעלייה ברווח הגנטי והפחתת הריבוי בשארות בבקר לחלב [Daetwyler 2007].

לרפתן בישראל, בנוסף לאימות הורות של פרות שבוצע להן שבב גנטי, ניתנים גם ערכי טיפוח גנומיים העוזרים בדירוג להוצאה של הפרה מהשירות. כללית חיזוי ערכי טיפוח גנומיים מוגבל בגלל גודל המדגם, אם כי אפשרי [Weller 2014]. כיום הרפתן מקבל את הערכים האלה במהלך האנליזה עם הפרים שמתבצעת בהולנד (CRV). הריצה כוללת imputation, אם צריך וחישוב ערכי הטיפוח הגנומיים תוך שילוב נתוני עדר הבקר ההולנדי [Weller 2015]. עד להיום נאספו כ- 9000 רשומות שבב גנטי עבור עגלות ופרות חולבות, כולל עתודות ופרות רגילות של משקים שאימצו טכנולוגיה זאת.

עבור כל תכונות הייצור, המתאם הגנטי בין פנוטיפים של פרים ופרות הוא כ- 0.6 והוא שונה משמעותית מ- 1 [Jenko 2017]. בעבר, גובשה שיטה לחישוב ישיר של ערכי טיפוח גנומיים והודגמה מהימנותה גם במקרה שהחישוב שלה מבוסס על אוכלוסיית ייחוס שבה המידע הפנוטיפי שונה בין תת-קבוצות, כלומר, עם אוכלוסיית ייחוס המורכבת גם מפרים וגם מפרות [Buch 2012]. עם זאת, בשיטה זאת, לא נלקחה בחשבון ההשפעה על המהימנות של ההטיה במדגם כלפי פרי הזרעה נבחרים. שילוב של פרות באוכלוסיית הייחוס יועיל מנקודת המבט של הקטנת ההטיה, משום

שבחירתן היא פחות אינטנסיבית. יש כמה פרסומים לגבי שימוש בפרות באוכלוסיית הייחוס עם נתונים אמיתיים [Ding 2013, Uemoto 2017, Pryce 2018]. אימוץ המסקנות מפרסומים אלה עשוי להוביל לבניית אוכלוסיית ייחוס מעורבת אמיתית, עבור עדר הבקר לחלב בישראל, בהתבסס על מודל דטרמיניסטי, שבו ניתן להגיע לתוצאות חד משמעיות, בתהליך אופטימיזציה. הטמעת מודל זה תשפר את המהימנות של ערכי טיפוח גנומיים החיוניים לתוכנית ההשבחה בישראל. בנוסף, אימות הורות, ומתן כלי מייעץ לחקלאי בדבר הפוטנציאל הגנטי של הפרה, ושל כדאיות הזיווג שלה בהתחשב בניבוי ריבוי בשארות גנומי (ROH, ראה תוצאות הקדמיות), יאפשרו שיפור מקומי בביצועי הרפתות ורווח כלכלי, שהערכתו המדויקת היא חלק ממטרות המחקר המוצע.

2. מטרות המחקר:

מטרת העל של המחקר המוצע היא להאיץ את קצב ההתקדמות הגנטית בתוכנית ההשבחה של בקר לחלב בישראל. מטרה זאת נקשרת למטרות הספציפיות הבאות:

2.1 התאמת מודל דטרמיניסטי לשילוב פרות באוכלוסיית הייחוס בחישוב ישיר של אומדני הורשה - Single step. הגדלת הגודל האפקטיבי של אוכלוסיית הייחוס, תאפשר הגדלת המהימנות והקטנת ההטיות של ערכי הטיפוח הגנומיים. זאת לצורך חישובם בדרך הנהוגה כיום וגם באלטרנטיבה הכרוכה בשימוש במערכות לומדות (תוכנית ענף: 362-0691, יישום מערכות לומדות (Machine Learning) לניבוי ערכי הטיפוח באמצעות סמנים של נוקלאוטיד יחיד (SNPs) באוכלוסיית בקר לחלב בישראל).

2.2 לבצע בקרת רשומות ההורות בספר העדר ולאתר ולתעד את הסיבות השונות לטעויות ושיעורן. בנוסף, לבחון חישוב ריבוי בשארות בהתבסס על הנתונים הגנומיים. על סמך רישומים אלה להעריך את הכדאיות הכלכלית ביישום כאמצעי להשבחה ולשיפור הממשק שנגזר מאימות הורות וניבוי מדויק יותר של ריבוי בשארות (איור 3).

2.3 להכשיר כוח אדם מתאים ליישום התחום ומחקרו.

2.4 להגדיל את המידע הגנטי של עדר הבקר לחלב בישראל על ידי ריצוף פרים מקווי דם שעדיין לא מיוצגים בבסיס הנתונים שבידינו ולהוסיף רשומות שבב גנטי לפרות רגילות מרפת וולקני שעדיין לא בוצע להן שבב.

3. חומרים, תהליכי ושיטות עבודה.

3.1 תוכנה וכוח אדם: גויס למחקר ד"ר אריה קורזון כמשתלם בת"ר דוקטור של האוניברסיטה העברית ובהדרכת חוקרי וולקני. למשתלם ניתנה גישה למחשב התאחדות מגדלי בקר והנחייה הדוקה עם מנהל ספר העדר, מר אפרים עזרא. לצורך התאמת מודל דטרמיניסטי לשילוב פרות באוכלוסיית הייחוס בחישוב ישיר - Single step, הותקנה תוכנת APEX- Linear Mixed Model Software (<https://ghpc.ai>) על תחנת עבודה (Ubuntu, ThinkStation P620) (256 GiB ram, 12 cores, 24 threads, Linux 20.04 LTS with AMD Ryzen™ Threadripper™ Processor) במכון לחקר בעלי חיים (וולקני, רחובות) לתקופת ניסיון של שנה.

APEX היא תוכנה עצמאית לניתוח נתונים כמותיים תוך שימוש במודלים מעורבים ליניאריים שעשויים להיות כרוכים בשונות שיתוף דטרמיניסטית בין רמות גורמים (למשל אילן יוחסין, אוטו-רגרסיבי), וסמנים גנומיים. בהתאם למודל הכמותי ולתנאי הרישוי, APEX מספקת פונקציונליות עבור:

- פתרון איטרטיבי של מערכת המשוואות,
- פתרון ישיר של מערכת המשוואות,
- הערכת רכיבי שונות מרובת משתנים באמצעות AI-REML או דגימת Gibbs מרובת שרשראות,
- חישוב שגיאות חיזוי מדויק ומהימנות,

• דגימת שגיאות לשונות ומהימנות,

• מודלים להערכת השגיאות הנ"ל.

ל- APEX אין מגבלות ביחס למספר המשוואות, למספר רמות הגורמים או למספר הגורמים המתואמים. הוא שימש לניתוח ריבוי תכונות עם יותר מ-1.3 מיליארד משוואות, יותר מ-40 מיליון פרטים באילן יוחסין, יותר מ-6 מיליון סמנים גנטיים ומעל 150 גורמים גנטיים בקורלציה

בעזרתו של הפרופסור האוסטרלי/דני- [Vinzent Böerner](#) הורצו בהצלחה דוגמאות לחישובים נדרשים על בסיס נתונים להדגמה. הוכן בסיס הנתונים הנדרש על בסיס ספר העדר בישראל. בוצעה בחינה של התאמת המודלים של APEX לשילוב פרות בחישוב ישיר - Single step, בתכנית ההשבחה בישראל.

3.2 יישום חישוב ישיר בחיזוי הערכות גנומיות לפרטים צעירים:

3.2.1 נתונים: מערך הנתונים כלל 9,763 פרטים גנוטיפיים, כולל 1,548 אבות ו-8,215 אימהות. ערכי טיפוח משוערים (EBV) עבור קילוגרמים של חלב, שומן וחלבון חושבו בחישוב ישיר של שלב יחיד (Curzon et al., 2024). עשרת הפרטים הצעירים ביותר עם תוצאות אמת בראשית 2025, הנחשבים כקשים ביותר לחיזוי, הוגדרו כקבוצת המבחן, והנתונים הגנטיים שלהם שולבו ללא ערכי הטיפוח משלב החישוב היחיד (Single-step genomic BLUP, ssGBLUP).

3.2. חיזוי ריבוי בשארות גנומי:

עבור חישוב ריבוי בשארות על ידי מקטעי (IBD) identity by descent, הותקנה תוכנת KING וחושבו מקטעי IBD1 ו-IBD2 כאשר השארות חושבה כחצי מפרופורציית הגנום שמקורה בהורה משותף. לחישוב שארות על ידי סימולציה של הורשת הפלוטיפים הותקנה תכנת Findhap לצורך שחזור סמנים ובניית הפלוטיפים בהורים. למידול הורשת הפלוטיפים הותקנה חבילת AlphaSimR ונערכה סימולציה של 100 צאצאים עם קצב רקומבינציה של 1 סנטימורגן ל 1Mb. חישוב run of homozygosity (ROH) בוצע בתכנת plink V1.90 ורמת השארות נקבעה בפרופורציית ההומוזיגוטיות שחושבה כממוצע ROH של 100 הסימולציות.

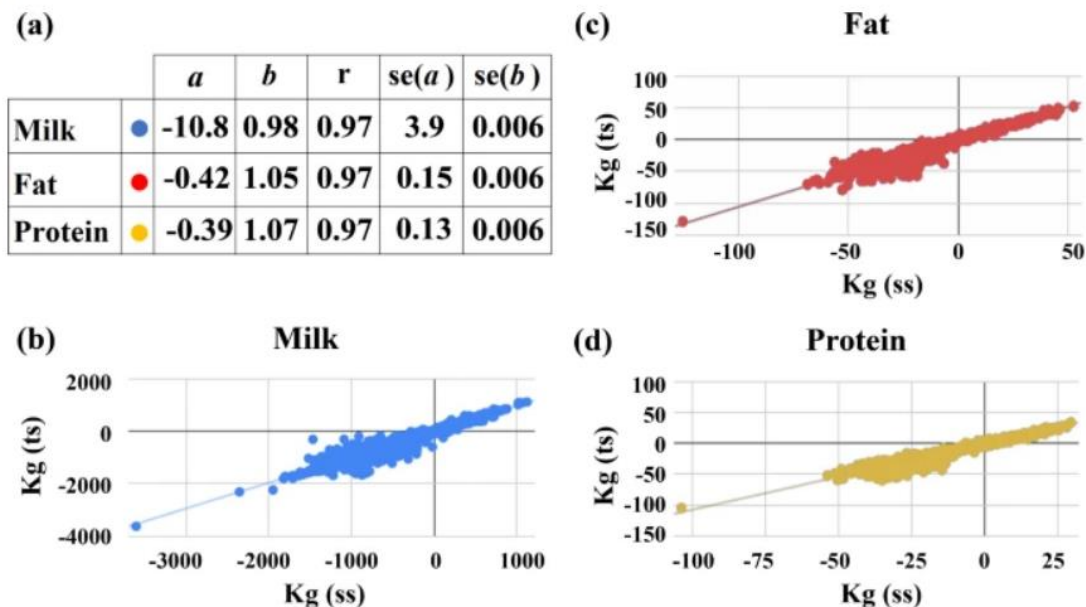
4. תוצאות.

יישום של מערכת הערכה גנומית יעילה דורש בנית אוכלוסיית ייחוס משמעותית. במערכת הקיימת (CRV) ממוזגים נתוני פרי ההזרעה מזן הולשטיין עבור שתי מדינות ישראל והולנד, כדי לקבל ערכי טיפוח גנומיים (gEBVs) בשיטה דו-שלבית מסורתית. הואיל שנתוני ספר העדר הנוכחיים כוללים אלפי שבבים גנטיים של פרות, תכנית מחקר זאת בוחנת שילוב של הנתונים הגנומיים האלה באוכלוסיית הייחוס. הערכת הפוטנציאל של הגנוטיפים של פרות נעשתה באמצעות יישום GBLUP בשלב יחיד (ssGBLUP) בעזרת תוכנת APEX. לשם הערכה זאת, השתמשנו בסט אימות שכלל נתוני הטיפוח משלב חישוב יחיד (ssGBLUP gEBVs) של 140 פרטים נוכחיים (2023, עם אמינות > 0.75), עבור שלוש תכונות ייצור (ק"ג חלב, ק"ג שומן וק"ג חלבון). עבור שלוש התכונות האלה, נתוני הטיפוח (gEBVs) החזויים של הפרטים על בסיס שנת 2019 ואילך, ללא מידע בנות, הושוו לערכי טיפוח גנומיים שחושבו עם מידע בנות משנת 2023 (טבלה 1). ערכי הטיפוח בטבלה זאת מופיעים כ- PTAs, שמשקפים את יכולת ההורשה כאשר $\text{Predicted Transmitting Ability} = \text{gEBVs}/2$. כמו כן, בהשוואה בין שיטות החד-ודו-שלבית (איור 1), הערכנו את השפעת המשקלים הפוליגנים, שמציינים את משקל מידע אילן היוחסין במטריצות היחסים הגנומיות (GRM, איור 2), וכן בדקנו את ההשפעה של השמטת רשומות ישנות וגנוטיפים ישנים.

דוח מסכם 362-0774 (יוני 2026)

עבור ייצור חלב, שומן וחלבון, נתוני הטיפוח משלב חישוב יחיד החזויים נמצאו בקורלציה חיובית עם ערכי טיפוח גנומיים שנצפו, בהתאם למשקלים הפוליגנים (ערכי z של 0.59, 0.66 ו-0.57, בהתאמה, על בסיס משקל פוליגני 0.5). לפיכך, איכויות התחזיות היו שוות לאלה של השיטה המסורתית הדו-שלבית של אוכלוסיה ממוזגת, שהן כיום 0.7, 0.56 ו-0.51, בהתאמה. באופן כללי, ערכי טיפוח גנומיים של שתי השיטות היו גם הם בקורלציה גבוהה ($r=0.97$, איור 1). גנוטיפים של 2.5, 6.5 ו-4.6 פרות בקירוב היו בעלי כוח סטטיסטי של פר יחיד עבור תכונות אלה, בהתאמה. השמטת רשומות נתונים ישנות מ-2000, שתרמה להגדלת מספר הגנוטיפים רק ב-1%, הובילה לחישוב ערכי טיפוח דומים לאלה שחושבו על בסיס שנת 2000 ואילך, בעוד שהשמטת גנוטיפים לפני 1995 השפיעה לטובה על החיזוי. מודל של חמש תחלובות חד-תכונתי היה עדיף על כל המודלים, כולל מודל ארבע תחלובות וארבע תחלובות רב-תכונתי. עבור כל המודלים, השוואה של ערכי טיפוח גנומיים חזויים ונצפים הצביעו על מתאמים גבוהים יותר, כאשר ערכי הטיפוח הגנומיים הוערכו ביתר (איור 2). הגדלת המשקל הפוליגני מתנה את הערכת היתר אך פגעה במתאמים (איור 2)

Fig. 1

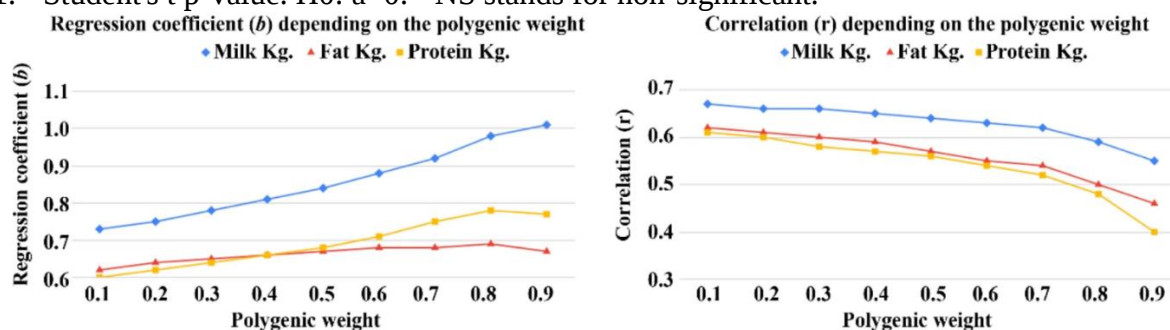


איור 1. השוואת ערכי PTA ע"פ שיטת ssGBLUP (ציר ה-X) שנבחנה בעבודה זו, עם ערכי PTA שחושבו בשיטת GBLUP ע"י CRV (ציר ה-Y). ערכי הרגרסיה ניתנים בפאנל (a) עבור קורלציה r , מקדם הרגרסיה, b , וחיתוך, a , לחלב (פאנל b), שומן (פאנל c) וחלבון (פאנל d), ניתנים.

טבלה 1. פרמטרי רגרסיה עבור 140 פרי אימות, בין ערכי טיפוח (PTAs) שחושבו מנתונים חלקיים למלאים.¹

Trait	Data	A ²	B	R ²	R	se(B)	se(a)	p-value B ³	p-value a ⁴
Milk (Kg)	Pedigree based BLUP (ABLUP)	7.37	0.91	0.16	0.4	0.18	77.56	NS ⁵	NS
	Data from 1990	-159.45	0.95	0.43	0.66	0.09	57.3	NS	0.006
	Data from 2000 (default model)	-157.92	0.96	0.43	0.66	0.09	57.64	NS	0.007
	Genotype animals from 1995	-141.73	0.99	0.42	0.65	0.1	55.17	NS	0.011
	Genotype animals from 2000	-157.92	0.96	0.43	0.66	0.09	57.64	NS	0.007
	Four-lactation model	-125.72	0.85	0.39	0.62	0.09	61.19	NS	0.042
	Four-lactation multi-trait model	-108.61	0.88	0.44	0.67	0.08	51.12	NS	0.035
	Current two-step model	-170.7	1.1	0.48	0.7	0.1	52.5	NS	0.001
Fat (Kg)	Pedigree based BLUP (ABLUP)	9.17	0.58	0.13	0.36	0.13	2.76	0.002	0.001
	Data from 1990	0.59	0.75	0.34	0.58	0.09	2.58	0.006	NS
	Data from 2000	0.58	0.75	0.34	0.59	0.09	2.52	0.006	NS
	Genotype animals from 1995	0.54	0.79	0.35	0.6	0.09	2.38	0.021	NS
	Genotype animals from 2000	0.58	0.75	0.34	0.59	0.09	2.52	0.006	NS
	Four-lactation model	2.17	0.67	0.31	0.56	0.08	2.64	0.00006	NS
	Four-lactation multi-trait model	3.31	0.7	0.33	0.57	0.09	2.53	0.001	NS
	Current two-step model	1.06	0.8	0.31	0.56	0.1	2.8	0.047	NS
protein (Kg)	Pedigree based BLUP (ABLUP)	10.23	0.41	0.03	0.18	0.19	2.77	0.002	0.0003
	Data from 1990	0.74	0.79	0.33	0.57	0.1	1.97	0.038	NS
	Data from 2000	0.22	0.81	0.32	0.57	0.1	2.03	NS	NS
	Genotype animals from 1995	0.33	0.86	0.31	0.56	0.11	2	NS	NS
	Genotype animals from 2000	0.22	0.81	0.32	0.57	0.1	2.03	NS	NS
	Four-lactation model	2.41	0.68	0.25	0.5	0.1	2.24	0.002	NS
	Four-lactation multi-trait model	2.61	0.69	0.26	0.51	0.1	2.24	0.002	NS
	Current two-step model	0.1	0.9	0.26	0.51	0.13	2.4	NS	NS

¹The truncated data was from 2000-2019 or 1990-2019 and the full data was from 2000-2023 or 1990-2023. For all models, the polygenic weights were set to 0.5. The current two-step model results are from the IBCA PTAs from 2019 and 2023. The two-step model contained 136 bulls that were included in the data from 2019 and 2023 and with reliability ≥ 0.75 in 2023. ²'a' is the regression intercept, B is the regression coefficient, R² is the coefficient of determination, r is the correlation and se stands for standard error. ³ Student's t p-value. H0: B=1. ⁴ Student's t p-value. H0: a=0. ⁵ NS stands for non-significant.



איור 2. קשר בין המשקל הפוליגני למקדמי הרגרסיה והמתאם בין התחזית לערך הנצפה. שינוי במשקל הפוליגני (משקל מטריצת הייחוסים הגנומית למטריצת השושלות) שיפר את הערכות היתר (מקדמי רגרסיה קרובים יותר ל-1, פאנל שמאל) אך פגע בדיוק (המתאמים, פאנל ימין).

תוצאות אלה פורסמו בכתב העת היוקרתי BMC genomics [Curzon 2024] בליווי סרטון תדמית (להלן קישורית):

[A new modeling approach could improve breeding for milk production in small dairy cattle populations](#)

דוח מסכם 362-0774 (יוני 2026)

ניתוח נוסף ניסה למתן את הערכות היתר תוך שימור, או שיפור המתאמים. במחקר זה נבחנו אסטרטגיות להפחתת הערכות היתר ושיפור ביצועי החיזוי בהערכת ssGBLUP עבור תנובת שומן וחלבון, באמצעות שילוב של ערכי טיפוח ממוכני-רגרסיה (Deregressed Proofs; DRP) ומידע בינלאומי של Interbull. בסיס הנתונים כלל עד 833,874 רשומות תנובת תחלובה ו-12,483 בעלי חיים בעלי מידע גנוטיפי. מודל ssGBLUP רב-תכונתי שיפר את הדיוק, אך עדיין הציג הערכות יתר (שיפועי רגרסיה של 0.67–0.75). לעומת זאת, מעבר למודל חד-תכונתי המבוסס על DRP עם שקלול אומדני מהימנות הפחית משמעותית את הערכות היתר, עם שיפועי רגרסיה שלא נבדלו באופן מובהק מ-1 (0.80–0.87, טבלה 2), תוך שמירה על דיוק דומה או מעט משופר.

טבלה 2. השוואה בין תחזיות ABLUP ו-ssGBLUP עבור תנובת חלבון ושומן עם משקל פוליגני קבוע של 0.5

Trait	Model	a	b	R ²	r	SE(b)	SE(a)	n
Fat	ABLUP	5.01	0.48	0.1	0.31	0.13	2.44	142
	ssGBLUP (5T)	-5.13	0.75	0.35	0.59	0.09	2.3	142
	ssGBLUP (DRP)	-3.25	0.87	0.34	0.58	0.1	2.17	142
	wssGBLUP (DRP)	-3.35	0.87	0.35	0.59	0.1	2.14	142
	ssGBLUP (DRP + IB bulls)	-3.06	0.88	0.33	0.58	0.1	2.17	142
	ssGBLUP (DRP + IB SNP)	-2.53	0.87	0.37	0.61	0.1	1.95	142
	ssGBLUP (DRP + IB bulls + SNP)	-2.30	0.87	0.37	0.61	0.1	1.95	142
	ABLUP	3.95	0.58	0.12	0.34	0.14	1.75	138
	ssGBLUP (5T)	-0.54	0.67	0.28	0.53	0.09	1.66	138
Protein	ssGBLUP (DRP)	0.47	0.81	0.27	0.52	0.11	1.56	138
	wssGBLUP (DRP)	0.47	0.80	0.28	0.53	0.11	1.55	138
	ssGBLUP (DRP + IB bulls)	0.48	0.82	0.27	0.52	0.12	1.57	138
	ssGBLUP (DRP + IB SNP)	1.14	0.80	0.28	0.53	0.11	1.44	138
	ssGBLUP (DRP + IB bulls + SNP)	1.14	0.81	0.28	0.53	0.11	1.45	138
	ssGBLUP (DRP + IB bulls + SNP) with outlier removal	1	0.84	0.30	0.56	0.12	1.42	137

השיפור הגדול ביותר בדיוק החיזוי התקבל באמצעות הכנסת משקלי סמנים שנגזרו מ-Interbull במסגרת פרויקט GenoEx, במיוחד עבור תנובת שומן, עם השפעה מזערית על הערכות היתר. באופן כללי, השימוש ב-DRP עם שקלול לפי מהימנות היה הגורם המרכזי בהפחתת הערכות היתר במודל ssGBLUP, בעוד שמידע בינלאומי, ובעיקר משקלי SNP, תרם בעיקר לשיפור דיוק החיזוי. ממצאים אלו מדגימים כי ניתן להתאים מודל עם שילוב מידע

חיצוני באופטימיזציה של הערכות גנומיות באוכלוסיות קטנות, דבר הפותח אפשרות לבסס הערכות גנומיות בישראל רק על סמך האוכלוסיה המקומית.

בשלב האחרון בדקנו האם חיצוי ריבוי בשארות על בסיס המטריצות הגנומיות יביא לשיפור בהערכות בהשוואה לשיטה הנהוגה כיום ומבוססת על תיעוד שושלות, וכן האן ניתן ליישמו כחלק מתכנית הטיפול. בעבר הראנו כי ניתן לשפר באופן מובהק את תחזית ריבוי בשארות על ידי שימוש בערכי ה- IBD של ההורים. כחלק מהצעה זו, בחנו מספר אסטרטגיות לדיוק וקלות ביישום.

מצאנו כי כל השיטות מבוססות הגנום הניבו תחזיות מדויקות יותר, עם מתאמים שנעו סביב 0.7, כאשר נראתה עדיפות קלה לשיטות מבוססות האפלוטיפ (כגון IBD וסימולציות הורשת הפלוטיפים משותפים, Roh_sim, טבלה 3). אך משיקולי פשטות ביישום השיטה, מצאנו כי שימוש במקדמי שארות המחושבים מתוך מטריצת GRM המחושבת כחלק ממודל ssGBLUP הנה הפשוטה ביותר ליישום במסגרת ההערכות הגנטיות. בנוסף, הראנו כי הגישות הגנומיות מאפשרות לבקר את השפעת כרומוזום X על ניבוי השארות, מה שמאפשר התאמת המודל לטיפול באמצעות זרמה ממוינת, בה קיים צפי ילודה של 90% נקבות.

טבלה 3. מתאמים בין ערכי ניבוי שארות לבין ערכי הומוזיגוטיות גנומיים המשקפים את הריבוי בשארות הנצפה. בשחור מודגש ערכי החיצוי מבוססי תיעוד שושלות הנהוגים המחושבים כיום.

Predictor	ROH5	ROH0.5	Diagonal Fgen
Fgen_base_pred	0.66	0.70	0.72
Fgen_0.5_pred	0.64	0.68	0.72
Roh_sim_5MB	0.66	0.67	—
Roh_sim_0.5MB*	0.68	0.72	—
IBD*	0.68	0.70	—
Fped	0.56	0.58	—

5. דיון ומסקנות.

עבור אוכלוסיות בקנה מידה קטן, כמו זאת של הבקר לחלב בישראל, גנוטיפים של פרות הנם בעלי ערך לבניית מערכי אימון בעלי כוח סטטיסטי משמעותי, המאפשרים חישוב של ערכי טיפוח גנומיים (gEBVs) אמינים. נתונים ישנים שכוללים גנוטיפים מועטים אינם משפרים את החיצוי. לעומת זאת, הגבלת השימוש בגנוטיפים לגנוטיפים עדכניים שיפרה את התחזיות של ערכי הטיפוח הגנומיים. כוונן הפרמטר של המשקל הפוליגני וטווח הזמן של דגימת הנתונים השפיע על דיוק של קביעת של ערכי הטיפוח הגנומיים והערכת יתר שלהם. לאור חשיבות הממצעים לתוכניות טיפוח במדינות בהן אוכלוסיית עדר החלב מוגבלת, כמו ישראל, הם סוכמו לידי מאמר, שפורסם בכתב העת BMC Genomics (קורזון וחובריו, 2024). שילוב תוכנת APEX ושיטת החישוב החד שלבי במערכת הטיפוח של התאחדות מגדלי בקר בישראל נדון בימים אלה, וזוכה לתמיכת מנהל ספר העדר.

במחקר המשך, הראנו כי שימוש בערכי DRP בשיקלול ערכי מהימנות והוספת משקלי סמנים מתוך האינטרבול ממתנים משמעותית את הערכות היתר תור שימור ואף שיפור דיוק תחזיות אומדני ההורשה. כל אלו מאפשרים ביצוע הערכות גנטיות בישראל ללא תלות בגופים חצוניים. כפי שהראנו, מעבר לשימוש בתחזיות גנומיות ויצור מטריצות ייחוסים גנומית, מאפשר גם שיפור בחיזוי ריבוי בשארות ותיקון תיעוד שושלות.

4. רשימת ספרות רלבנטית (* כוכביות מסמנות מובאות בהן שותפים הכותבים):

Alkhoder H, Liu Z, Segelke D, Reents R. 2022 Comparison of a single-step with a multistep single nucleotide polymorphism best linear unbiased predictor model for genomic evaluation of conformation traits in German Holsteins. J Dairy Sci. 105:3306-3322.

Buch LH, Kargo M, Berg P, et al. 2012 The value of cows in reference populations for genomic selection of new functional traits. Animal. 6:880-6.

Calus MPL, de Haas Y, Veerkamp RF. 2013 Combining cow and bull reference populations to increase accuracy of genomic prediction and genome-wide association studies. J Dairy Sci. 96:6703-15.

Cohen-Zinder M, Shor-Shimoni E, Lipkin E, Ben-Meir Y, Asher A, Miron J, Shabtay Ariel. 2019 *FABP4* gene has a very large effect on feed efficiency in lactating Israeli Holstein cows. Physiological Genomics. 51:481-487.

*Curzon AY, Ezra E, Weller JI, **Seroussi E**; Borner V, **Gershoni M**. 2024 Single-step genomic BLUP (ssGBLUP) effectively models small cattle populations: lessons from the Israeli-Holstein Herdbook. BMC Genomics 25: 1147.

Daetwyler HD, Villanueva B, Bijma P, et al. 2007 Inbreeding in genome-wide selection. J Anim Breed Genet. 124:369-76.

de Oliveira HR, Brito LF, Sargolzaei M, e Silva FF, Jamrozik J, DAL Lourenco, FS Schenkel. 2019 Impact of including information from bulls and their daughters in the training population of multiple-step genomic evaluations in dairy cattle: A simulation study. Journal of Animal Breeding and Genetics, 136:441-452.

Ding X, Zhang Z, Li X, et al. Accuracy of genomic prediction for milk production traits in the Chinese Holstein population using a reference population consisting of cows. J Dairy Sci. 2013;96:5315-23.

***Gershoni M**, Shirak A, Raz R, **Seroussi E**. 2022 Comparing BeadChip and WGS Genotyping: Non-technical failed calling is attributable to additional variation within the probe target sequence Genes. 13, 485.

Falconer DS, Mackay TFC. 1996 Introduction to quantitative genetics. 4th ed. Harlow, UK: Longman.

Freeman AE. 1991 CR Henderson: Contributions to the dairy industry. J Dairy Sci. 74:4045-4051.

Jenko J, Wiggans GR, Cooper TA, et al. 2017 Cow genotyping strategies for genomic selection in a small dairy cattle population. J Dairy Sci. 100:439-52.

Lourenco D, Legarra A, Tsuruta S, Masuda Y, Aguilar I, Misztal I. 2020 Single-step genomic evaluations from theory to practice: using SNP chips and sequence data in BLUPF90. Genes. 11:790.

Koivula M, Strandén I, Aamand GP, Mäntysaari EA. 2016 Effect of cow reference group on validation reliability of genomic evaluation. Animal. 10:1061-1066.

- Meuwissen THE, Hayes BJ, Goddard ME. 2007 Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157:1819–29.
- Mäntysaari E, Liu Z, vanRaden PM. 2010 Interbull validation test for genomic evaluations. *Interbull bulletin*. 41:4–5.
- Misztal I, Legarra A, Aguilar I. 2009 Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J J Dairy Sci*. 92: 4648-4655.
- Pryce JE, Nguyen TTT, Axford M, et al. 2018 Symposium review: Building a better cow—the Australian experience and future perspectives. *J Dairy Sci*. 101:3702–13.
- Ron M, Blanc Y, Band M, Ezra E, Weller JI. 1996 Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement, *J Dairy Sci*. 79:676-81.
- Schaeffer LR. 2006 Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet*. 123:218–23.
- Uemoto Y, Osawa T, Saburi J. 2017 Effect of genotyped cows in the reference population on the genomic evaluation of Holstein cattle. *Animal*. 11:382–93.
- VanRaden PM. (2008). Efficient methods to compute genomic predictions. *J Dairy Sci*. 91:4414-4423.
- VanRaden PM. 2017. Genomic tools to improve progress and preserve variation for future generations. EAAP abstract 26496 and slide presentation (<https://slideplayer.com/slide/12943174/>).
- Wiggans GR, Cooper TA, VanRaden PM, Cole JB. 2011 Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. *J Dairy Sci*. 94: 6188-6193.
- *Weller JI, Glick G, Shirak A, Ezra E, **Seroussi E**, Shemesh M, Zeron Y, Ron M. 2014 Predictive ability of selected subsets of single nucleotide polymorphisms (SNPs) in a moderately sized dairy cattle population. *Animal* 8:208–16.
- Weller JI, Stoop WM, Eding H, Schrooten C, Ezra E. 2015 Genomic evaluation of a relatively small dairy cattle population by combination with a larger population. *J Dairy Sci*. 98:4945–55.