

DOKTORI (PhD) ÉRTEKEZÉS

BOGNÁR LÁSZLÓ

MOSONMAGYARÓVÁR

2025

DOKTORI (PhD) ÉRTEKEZÉS

**SZÉCHENYI ISTVÁN EGYETEM
ALBERT KÁZMÉR MOSONMAGYARÓVÁRI KAR
ÁLLATTUDOMÁNYI TANSZÉK**

**WITTMANN ANTAL NÖVÉNY-, ÁLLAT- ÉS ÉLELMISZERTUDOMÁNYI
MULTIDISZCIPLINÁRIS DOKTORI ISKOLA**

UJHELYI IMRE ÁLLATTUDOMÁNYI DOKTORI PROGRAM

DOKTORI ISKOLA VEZETŐ:

DR. VARGA LÁSZLÓ, AZ MTA DOKTORA, EGYETEMI TANÁR

DOKTORI PROGRAMVEZETŐ:

DR. SZABÓ FERENC, AZ MTA DOKTORA, PROFESSOR EMERITUS

TUDOMÁNYOS VEZETŐK:

DR. SZABÓ FERENC, AZ MTA DOKTORA, PROFESSOR EMERITUS

DR. ANTON ISTVÁN, AZ MTA DOKTORA, PROFESSOR EMERITUS

ÉRTEKEZÉS CÍME:

**A GENOMIKAI TENYÉSZÉRTÉKBECSLÉS ÉS A GENOMSZELEKCIÓ
MAGYARORSZÁGI BEVEZETÉSÉNEK GYAKORLATI
TAPASZTALATAI A HOLSTEIN-FRÍZ FAJTA ESETÉBEN**

KÉSZÍTETTE:

BOGNÁR LÁSZLÓ

MOSONMAGYARÓVÁR

2025

**A GENOMIKAI TENYÉSZÉRTÉKBECSLÉS ÉS A GENOMSZELEKCIÓ
MAGYARORSZÁGI BEVEZETÉSÉNEK GYAKORLATI
TAPASZTALATAI A HOLSTEIN-FRÍZ FAJTA ESETÉBEN**

Írta:

BOGNÁR LÁSZLÓ

**Készült a Széchenyi István Egyetem Wittmann Antal Növény-, Állat- és
Élelmiszertudományi Multidiszciplináris Doktori Iskola
Ujhelyi Imre Állattudományi Doktori Programja keretében**

Témavezető: Dr. Szabó Ferenc, az MTA doktora, professor emeritus

Elfogadásra javaslom (igen / nem)

(aláírás)

A jelölt a doktori komplex vizsgán megfelelt.

Mosonmagyaróvár,

.....
a Komplex Vizsga Bizottság elnöke

Az értekezést bírálóként elfogadásra javaslom (igen/nem)

Első bíráló: Dr. Wagenhoffer Zsombor, PhD, intézetvezető (ÁOE) igen/nem

(aláírás)

**Második bíráló: Dr. habil. Mikó Edit, egyetemi docens, dékán (SZTE-MGK)
igen/nem**

(aláírás)

**Esetleg harmadik bíráló (Dr.)
igen/nem**

(aláírás)

A jelölt az értekezés nyilvános vitáján%-ot ért el.

Mosonmagyaróvár,

A Bírálóbizottság elnöke

Doktori (PhD) oklevél minősítése.....

Az EDT elnöke

A rövid értekezés háttere

„A Wittmann Antal Növény-, Állat- és Élelmiszertudományi Multidiszciplináris Doktori Iskola működési szabályzatának 1. sz. melléklete alapján a „klasszikus” doktori értekezés helyett lehetőség van ún. rövid értekezés készítésére. Ebben a doktorjelölt tömören foglalja össze a doktori kutatómunkája témakörében megjelent tudományos közleményeiben publikált eredményeit, mellékelve magukat a közleményeket is. A rövid értekezés és a mellékelt publikációk együtt képezik a doktori művet. A rövid értekezésnek elsőszerzős, a **SCImago** adatbázisában (<https://www.scimagojr.com/journalrank.php>) szereplő Q1–Q3 besorolású folyóiratokban megjelent, illetve megjelenésre elfogadott, teljes terjedelmű, eredeti közleményeken (full-length original research paper) kell alapulnia. E közleményeknek összesen minimum 6 pontot kell érniük az alábbi számítás szerint úgy, hogy közülük legalább az egyiknek Q1-es vagy Q2-es folyóiratban megjelentnek (vagy megjelenésre elfogadottnak) kell lennie. Q1-es cikk: 4 pont, Q2-es cikk 2 pont, Q3-as cikk: 1 pont.

A feltételek teljesülése a csatolt publikációk alapján:

<i>Animals</i> (2024) 15 (51),11. https://doi.org/10.3390/ani15010051 (Q1)	4 pont
Different Breeding Values Under Uniform Environmental Condition for Milk Production Yield Traits in Holstein-Friesian Cows	
<i>Animals</i> (2024) 14 (23), 3518. https://doi.org/10.3390/ani14233518 (Q1)	4 pont
Simultaneous Effects of Single-Nucleotide Polymorphisms on the Estimated Breeding Value for Milk, Fat, and Protein Yield of Holstein-Friesian Cows in Hungary	
<i>CET</i> (2023) 107, 169–174. https://www.cetjournal.it/cet/23/107/029.pdf (Q3)	1 pont
Management of “Modern” Holstein Cows Focusing on Sustainability and Resilience	
Összesen:	9 pont

TARTALOMJEGYZÉK

AZ ÉRTEKEZÉSBEN TALÁLHATÓ SZAKKIFEJEZÉSEK MAGYARÁZATA	7
KIVONAT	9
ABSTRACT	10
1. BEVEZETÉS ÉS CÉLKITŰZÉSEK	11
2. ANYAG ÉS MÓDSZER	13
2.1. A genomanalízis és a genomikai tenyésztérbecslés módszere	13
2.2. BLUP-tenyésztérbecslési módszerek összehasonlítása	16
3. EREDMÉNYEK ÉS ÉRTÉKELÉSÜK	18
3.1. SNP-vizsgálatok eredménye	18
3.2. BLUP-tenyésztérbecslési módszerek összehasonlító eredménye	20
4. ÚJ TUDOMÁNYOS EREDMÉNYEK ÉS HASZNOSÍTÁSUK	24
4.1. Új tudományos eredmények	24
4.2. Az eredmények hasznosíthatósága	25
5. AZ ÉRTEKEZÉS TÉMAKÖRÉBEN MEGJELENT PUBLIKÁCIÓK	26
5.1. Idegen nyelvű folyóiratban megjelent lektorált cikk	26
5.2. Magyar nyelvű folyóiratban megjelent lektorált cikk	26
6. MELLÉKLETEK	27
6.1. Melléklet – <i>B. taurus</i> genom ARS-UCD1.2	27
6.2. Melléklet – Az egyes markerekhez kapcsolódó vizsgálati értékek	31
6.3. Melléklet – a négy kandidáns gén jellemzői	33

AZ ÉRTEKEZÉSBEN TALÁLHATÓ SZAKKIFEJEZÉSEK MAGYARÁZATA

Allél: alternatív gén- vagy szekvenciaváltozat a kromoszóma adott helyén.

BLUP (Best Linear Unbiased Prediction): a legjobb lineáris torzítatlan előrejelzés.

Chip (SNP): egy microarray-technológiai eszköz, amelyet genetikai variációk, például SNP-k nagy áteresztőképességű vizsgálatára használnak. Főként genotipizálásra alkalmazzák az állattenyésztésben, a növénynevelésben és az emberi genetikai kutatásokban.

CRV (Arnhem, Hollandia): holland és flamand gazdák által tulajdonolt szövetkezet, amely innovatív genetikai megoldásokkal, szaporítóanyag-szolgáltatásokkal és fejlett adatmenedzsment-rendszerekkel segíti a szarvasmarha-tenyésztés hatékonyságát, a termelés fenntarthatóságát és az állomány genetikai potenciáljának optimalizálását.

DGV (Direct Genomic Value): direkt genomikai érték, a tenyésztérbecslésben egy állat genetikai értékének közvetlen, genomialis adatok alapján történő becslését jelenti. Ez az érték kizárólag genotípusinformációkon alapul, és nem veszi figyelembe az állat saját fenotípusát vagy származási (pedigré-) információit.

DNS-marker: olyan genomialis hely, lókus, ahol a populációra vonatkoztatva kettő vagy több változat, allél található meg. A DNS-markerek leginkább elterjedt változatai az SNP-k.

DRP (Deregressed Proof): a hagyományos tenyésztérből származtatott de-regresszált tenyésztérb.

EuroGenomics: Az EuroGenomics egy nemzetközi együttműködésen alapuló konzorcium, amelynek célja a szarvasmarha-tenyésztés genomikai alapú fejlesztése. A szervezet Európa vezető szarvasmarha-tenyésztő szövetségeit és kutatóintézeteit fogja össze, hogy megosszák genetikai erőforrásaikat és szakértelmüket. Az EuroGenomics által fejlesztett SNP-chip technológia, például az EuroG_MDv4, nagy felbontású genotipizálást biztosít, amely kulcsfontosságú a tenyésztérbecslésben és a genomikai szelekcióban, lehetővé téve a gazdaságilag fontos tulajdonságok gyorsabb és pontosabb genetikai javítását.

FDR: (False Discovery Rate): téves azonosítási ráta vagy hibás felderítési arány. Annak a valószínűsége, hogy rossz eredményt fogadunk el (false positive).

Fenotípus: az adott egyed külső vagy mérhető tulajdonságai, illetve tulajdonságainak összessége.

Fst-marker: az SNP-k közötti genetikai távolság.

GE BV (Genomically Enhanced Breeding Value): genomikai információkkal bővített tenyésztérb.

Genotípus: az egyed egy vagy több DNS-lókusán meghatározott marker típusa, hossza vagy az ott elhelyezkedő nukleotidok megnevezése.

Lókus: a DNS-lánc egyedi, azonosítható helye.

MACE (Multiple Across Country Evaluation): határokon átívelő többszörös nemzetközi tenyésztérbécslés (lásd <https://interbull.org/ib/interbullactivities>).

Nukleotid: A DNS és az RNS alapvető építőkőve, egy olyan összetett molekula, amely egy nitrogéntartalmú bázisból, egy cukorból és egy foszfátcsoportból áll.

Populáció: egy adott faj azon egyedeinek csoportja, akik egy meghatározott földrajzi területen élnek, képesek egymással szaporodni, és közös génállománnyal (génkészlettel) rendelkeznek. A populáció tagjai közötti genetikai és ökológiai kölcsönhatások alakítják a csoport dinamikáját, genetikai sokféleségét és alkalmazkodóképességét.

QTL (Quantitative Trait Locus): régió (lókusz) a kromoszómán, ahol a mennyiségi tulajdonságok kialakításában szerepet játszó gének helyezkednek el.

SNP (Single Nucleotide Polymorphism): egy pontos nukleotidpolimorfizmus a DNS-ben található genetikai variáció egyetlen nukleotid (A, T, C vagy G) szintjén bekövetkező eltérését jelenti egy populáció egyedei között. Ez a genetikai variáció a genomialis különbségek egyik leggyakoribb formája, amely fontos szerepet játszik a genetikai kutatásokban és alkalmazásokban. Akkor tekintünk egy variációt SNP-nek, ha a populáció legalább 1%-ában megjelenik.

SNV (Single Nucleotide Variance): egyetlen nukleotidbázist érintő eltérés a populáció egyedeiben.

SSVS (Stochastic Search Variable Selection): sztochasztikus keresési változó kiválasztása.

KIVONAT

A tanulmány célja a holstein-fríz (HF) tehének tej-, zsír- és fehérjetermeléséhez kapcsolódó tenyésztékek genetikai hátterének vizsgálata és validálása volt. A Holstein-fríz Tenyésztők Egyesülete (HFTE) által működtetett HUNGENOM program keretében genotipizálási vizsgálatokat végeztünk EuroGenomics (EuroG_MDv4) chip alkalmazásával. A tejhozam (TH), zsírhozam (ZSH) és fehérjehozam (FH) tenyésztékadatait a HFTE biztosította. A tenyésztékekkel összefüggő lókuszokat három módszerrel azonosítottuk: (1) az SNP-k genetikai távolságának (F_{st_marker}) meghatározásával, (2) lineáris regressziós elemzéssel és (3) haplotípus-asszociációs tesztekkel. Az azonosított SNP-k közül több olyan gén közelében helyezkedik el, amelyek a szarvasmarhák esetében még nem kerültek vizsgálatra, de amelyek a tejtermelési tulajdonságokhoz kapcsolódó jövőbeli kutatások kiemelt célpontjai lehetnek. Eredményeink rávilágítottak arra, hogy a két vagy három tenyésztékekkel összefüggő markerek hatékonyan alkalmazhatók a szelekcióban, elősegítve a genetikai előrehaladás felgyorsítását.

Ezek az információk különösen hasznosak lehetnek a tenyésztési és szelekciós célkitűzések elérésében. Ugyanakkor az eredmények azt is megmutatták, hogy bár a legtöbb marker hatása a tenyésztékek irányultságával összhangban van, néhány marker alkalmazása óvatosságot igényel, mivel azok eltérő hatást gyakorolhatnak az egyes tenyésztékekre – például miközben növelik az egyiket, addig egy vagy akár mindkét másik tenyészték csökkenését eredményezhetik.

A tanulmány második részében a HF tehének fenotípusos teljesítményét és tenyésztékét vizsgáltuk, összehasonlítva három eltérő tenyésztékbecslési megközelítést: a pedigrén alapuló BLUP-ot, a hagyományos BLUP-ot, valamint a genomikai BLUP-ot. Az összehasonlítás eredményei alapján a hagyományos BLUP bizonyult a legmegbízhatóbb módszernek, amelyet a genomikai adatokkal kiegészített BLUP követett. A pedigrén alapuló BLUP-módszer megbízhatósága volt a leggyengébb.

A tanulmány újdonsága abban rejlik, hogy a tej-, zsír- és fehérjetermelésre vonatkozó tenyészték-előrejelzéseket úgy validáltuk, hogy a különböző módszerekkel készült becsléseket közvetlenül összehasonlítottuk a tényleges fenotípusos teljesítményekkel, azonos telep, év és évszak környezeti feltételek mellett. Az eredmények jelentős hozzájárulást nyújtanak a genomikai szelekció alkalmazásának továbbfejlesztéséhez és a tenyésztékbecslés pontosságának növeléséhez.

ABSTRACT

The study aimed to investigate and validate the genetic basis of breeding values associated with milk, fat, and protein production in Holstein-Friesian (HF) cows. Genotyping was conducted as part of the HUNGENOM program led by the Hungarian Holstein-Friesian Breeders Association (HFTE), using the EuroGenomics (EuroG_MDv4) chip. Data on breeding values for milk yield (MY), fat yield (FY), and protein yield (PY) were provided by the HFTE. Genomic loci associated with the breeding values were identified using three methods: (1) calculation of SNP genetic distances (F_{st_marker}), (2) linear regression analysis, and (3) haplotype association tests. Several identified SNPs were located near genes that have not yet been studied in cattle but may serve as critical targets for future research related to production traits. Our results highlighted that markers associated with two or three breeding values can be effectively utilized in selection programs to accelerate genetic progress.

This information is particularly valuable for achieving breeding and selection objectives. However, the findings also revealed that while most markers had consistent effects across breeding values, some require cautious application due to their contrasting effects – for instance, increasing one breeding value while decreasing another or even both.

In the second part of the study, the phenotypic performance and breeding values of HF cows were analysed by comparing three distinct breeding value estimation methods: pedigree-based BLUP, traditional BLUP, and genomic BLUP. The results showed that traditional BLUP provided the most reliable estimates, followed by genomic BLUP. Pedigree-based BLUP exhibited the lowest reliability among the three methods.

The novelty of this study lies in its validation of breeding value predictions for milk, fat, and protein production by directly comparing estimates derived from different methods with actual phenotypic performance under consistent herd, year, and season conditions. These findings provide valuable insights into the practical application of genomic selection and contribute to improving the accuracy of breeding value estimation.

1. BEVEZETÉS ÉS CÉLKITŰZÉSEK

A holstein-fríz (HF) világszerte a legismertebb és leghatározottabban tejelő szarvasmarhafajta, amely kiemelkedő tejtermelő képességgel rendelkezik. Az elmúlt évek során a tenyészkiválasztás, nemesítés és szelekció alapelvei változatlanok maradtak: a legjobb szülők párosításával az utódok tulajdonságait igyekszünk javítani. Ugyanakkor az információ, amelyen a szelekciós döntések alapulnak, egyre pontosabbá vált. Korábban elsősorban a fenotípusos teljesítmények alapján rangsorolták az állatokat, különös tekintettel a tenyész bikák lányainak tejtermelési eredményeire. A fenotípusos teljesítményt a $Fenotípus = Genotípus \times Környezet$ ($P = G \times E$) képlet írja le, amely azt mutatja, hogy a fenotípus a genetikai adottságok és a környezeti tényezők kölcsönhatásából adódik. A környezeti tényezők – mint a tartási körülmények, takarmányozás és menedzsment – az elért teljesítmény 60–70%-át befolyásolják, míg a genetikai alapok 30–40%-ot tesznek ki. A nemesítés célja, hogy különböző módszerek alkalmazásával minél pontosabban megbecsüljük az egyedek genetikai adottságait, és ezek ismeretében célzott párosításokat hajtsunk végre. Noha a környezeti tényezők optimalizálhatók, a tenyészcélok eléréséhez elsősorban azok a tenyészállatok járulnak hozzá, amelyek azonos tartási körülmények között is képesek nagyobb teljesítményre és hatékonyabb termelésre. A tehénállomány koncentrációjával párhuzamosan a genomika és a biotechnológia robbanásszerű fejlődésének is tanúi lehettünk. A Humán Genom Projekt tudományos áttöréseit követően néhány éven belül szinte minden gazdasági szempontból jelentős állatfaj és fajta génszekvenálása megtörtént. Ennek eredményeként a genotipizálásból származó adatok mennyisége exponenciálisan növekedett, és idővel lehetővé vált azok felhasználása populációgenetikai összefüggések feltárására.

Az egyedi örökítőanyagban megfigyelhető különbségek, amelyek bázispárok szintjén pontszerű eltéréseként, azaz nukleotidvarianciaként (korábbi szóhasználattal mutációként) jelentkeznek, alapvető jelentőséggel bírnak a genetikai kutatásokban. Amennyiben egy adott variáns előfordulása a populációban eléri az 1%-ot, polimorfizmusról beszélünk (Single Nucleotide Polymorphism, SNP). Ezek a polimorfizmusok megfelelő statisztikai módszerek alkalmazásával összefüggésbe hozhatók fenotípusos tulajdonságokkal, például a tejtermeléssel, funkcionális küllemmel, vagy egyéb életfolyamatokhoz és menedzsmenthez kapcsolódó jellemzőkkel. A genotipizálás során feltárt SNP-k, a hagyományos módszerekkel (pl. BLUP) becsült tenyészértékek és a fenotípusok összefüggő rendszere lehetővé teszi az egyedek teljesítményének nagy megbízhatóságú előrejelzését. Ez különösen fontos a korai tenyész kiválasztásban, mivel a saját teljesítménnyel még nem rendelkező egyedek, például fiatal tenyészállatok és tenyész bikajelöltek későbbi életszakaszban realizálható termelési potenciálja jól prognosztizálható.

A genomszelekció alapfeltétele a referenciapopuláció, más néven tréningpopuláció megléte és folyamatos naprakészen tartása. A referenciapopuláció fenntartása nélkülözhetetlen a pontos genomikai tenyésztérbecslés érdekében. A tréningpopuláció olyan egyedekből áll, amelyek esetében a saját teljesítményeik, valamint származási információik alapján becsült (hagyományos) tenyésztérbük rendelkezésre áll, továbbá ismert az egyedi SNP-térképük is. Ez utóbbi az örökítőanyaguk (DNS) alkalmazott vizsgálati platform által meghatározott helyeinek pontszerű genetikai változásait tartalmazza. A szarvasmarha-tenyésztésben az Illumina, Inc. (San Diego, CA, USA) cég 50k-s chipjének használata az általánosan elterjedt, amely 50 000, meghatározott helyeken található pontszerű mutáció előfordulását/állapotát vizsgálja az DNS-ben. Könnyen belátható, hogy kellő méretű referenciapopuláció esetében pusztán az SNP-információ alapján az adott egyed teljesítménye nagy megbízhatósággal előre jelezhető. Ez a HF fajta esetében 3 000, mindkét típusú tenyésztérbükkel rendelkező tenyészbikát tartalmazó referenciapopuláció használatával a becslés megbízhatósági értéke (Rel.) megközelíti a 70%-ot. A genomikai tenyésztérbükbecslés eredményeire alapozott nemesítő munka a fenti összefüggések következtében rendkívül eredményes, és többek között ez magyarázza robbanásszerű elterjedését. A módszer további előnye a generációs intervallum csökkentése, hiszen a születést követő legkorábbi életszakaszban vett biológiai minta, illetve a korszerű biotechnikai módszerek alkalmazásával akár az embriókból biopsziával nyert néhány sejtből kivont DNS elemzésével is teljes és megbízható tenyésztérbük-információt kaphatunk a még meg sem született egyed képességeiről és hozhatunk megalapozott tenyésztési döntéseket.

Az előbb említett szempontokat alapul véve az alábbi vizsgálati célokat tűztem ki:

1. a hazai HF populáció genomikai tenyésztérbükbecslésének bevezetése, illetve a kapott tenyésztérbük populációs szinten történő validálása,
2. első laktációs HF tehenek termelési tulajdonságainak tenyésztérbükbecslése három különböző módszerrel (pedigrén alapuló, hagyományos és genomikai BLUP),
3. a három BLUP-módszer pontosságának, előrejelző képességének, valamint megbízhatóságának értékelése azonos évben, évszakban, illetve ugyanabban az állományban született és tartott első laktációs tehenek esetében,
4. az egyes becslési eljárások alkalmazhatóságának vizsgálata a modern, precíziós tenyésztési programokban,
5. az SNP-chip segítségével tipizált azon SNP-k meghatározása, amelyek egyszerre két, illetve három tulajdonságra becsült TÉ kialakításában is szerepet játszanak,
6. a fenti szempontok alapján kiválasztott SNP-k és a TÉ-ek regressziós együtthatóinak (a változások irányának és mértékének) meghatározása.

2. ANYAG ÉS MÓDSZER

2.1. A genomanalízis és a genomikai tenyésztértékbecslés módszere

A vizsgálatot a HUNGENOM projekt keretében végeztük, melynek során az Illumina cég EuroG MD BeadChip-jét használtuk. Kiterjedt nemzetközi összefogás eredményeként a magyar mintákat feldolgozó laboratórium hozzáférést biztosított számunkra az SNP-adatok milliárdjaihoz, amelyen a tenyésztértékbecslés és saját vizsgálatunk alapult. A DNS elemzéséből származó, egyedi mutációváltozatokat leíró SNP-mintázatot a már említett referencia-adatbázis értékeivel vetettük össze annak érdekében, hogy megbízható egyedi genomikai tenyésztértékeket kapjunk.

A saját vizsgálat genomanalízis részében 2 963 tehén adata szerepelt. Az állatok fenotípusára és genotípusára vonatkozó valamennyi adatot a Holstein-fríz Tenyésztők Egyesülete (HFTE) bocsátotta rendelkezésre. A HF tehének genotipizálását a 67 227 SNP-t tartalmazó EuroG_MDv4 microarray-vizsgálat (EuroGenomics, Hollandia, Amszterdam) segítségével végeztük. Csak a 0,95-nél magasabb találati aránnyal (call rate) rendelkező mintákat vettük figyelembe ebben a vizsgálatban. A végleges adathalmaz így 59 151 SNP-t tartalmazott. A tenyésztértékeket a HFTE adataira fejlesztett tenyésztértékbecslési modellel (CRV Arnheim, Hollandia) határoztuk meg, amely az X kromoszómán található SNP-k kivételével 40 947 SNP-t használt fel.

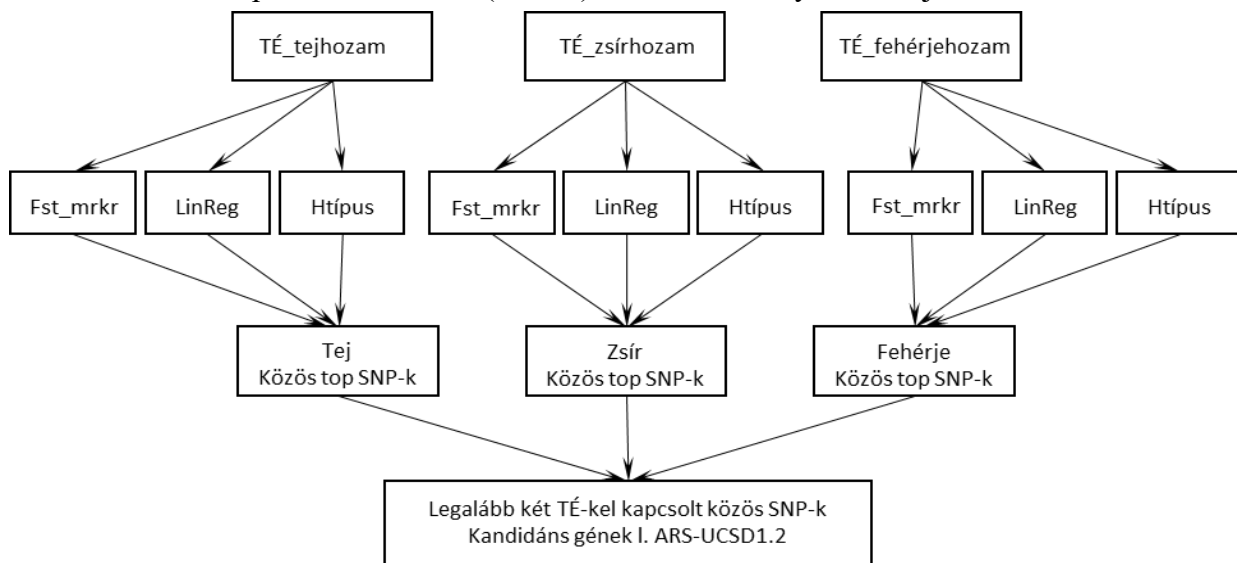
A genomikai becslési modell a Bayes-féle multi-QTL modellen alapul, ahol az SNP-k hatását közvetlenül értékelik, haplotípusok vagy származási adatok használata nélkül. Bár a módszer egyszerre több tulajdonságra is alkalmazható, a rutinszerű genomikai értékelések egy tulajdonságra vonatkozó elemzések, azaz $m = 1$. M tulajdonságra a modell a következő:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{u}_i + \sum_{j=1}^{40947} \mathbf{z}_{ij} \mathbf{q}_j \mathbf{v}_j + \mathbf{e}_i,$$

ahol \mathbf{y}_i ($m \times 1$) vektor, az i . bika fenotípusainak (DRP) vektora, $\boldsymbol{\mu}$ ($m \times 1$) vektor az adott tulajdonság fix átlaga, \mathbf{u}_i ($m \times 1$) vektor, az i . bikához tartozó random poligén hatás, \mathbf{q}_j (3×1) vektor, a j . SNP random, skálázatlan hatása, amely a 0., az 1. és a 2. allélhoz tartozik (a 0. allél a hiányzó genotípus-információnak felel meg), \mathbf{v}_j ($1 \times m$) vektor, a j . SNP-hez tartozó random skálafaktorok, \mathbf{e}_i ($m \times 1$) vektor, az i . bikához tartozó reziduum (a modell hibája), és \mathbf{z}_{ij} az i . bika j . SNP-jéhez tartozó designvektor. $\mathbf{z}_{ij} = [0 \ 2 \ 0]$ és $[0 \ 0 \ 2]$ homozigóta (AA, illetve BB) bikák, $\mathbf{z}_{ij} = [0 \ 1 \ 1]$ heterozigóta (AB) bikák, és $\mathbf{z}_{ij} = [2 \ 0 \ 0]$ azon bikák esetén, ahol az adott SNP-hez tartozó információk hiányoznak.

A GWAS- (teljes genomra kiterjedő asszociációs) vizsgálat során a teheneket a tejhozam (TH), zsírhozam (ZSH) és fehérjehozam (FH) tenyésztértékük ($TÉ_{tej}$, $TÉ_{zsír}$, $TÉ_{fehérje}$) szerint osztottuk fel. Először minden $TÉ$ -kategóriában magas és alacsony értékű csoportokat hoztunk létre a következők szerint: $TÉ_{tej_magas} > 1\,465$, $TÉ_{tej_alacsony} < 328$; $TÉ_{zsír_magas} > 65$, $TÉ_{zsír_alacsony} < 19$; $TÉ_{fehérje_magas} > 51$, $TÉ_{fehérje_alacsony} < 21$. Minden tulajdonságot ($TÉ_{tej}$, $TÉ_{zsír}$, $TÉ_{fehérje}$) három független megközelítéssel határoztunk meg: az SNP-k genetikai távolsága (Fst_marker), a lineáris regresszió és a haplotípus-asszociációs teszt segítségével, az SNP and Variation Suite (SVS) szoftver felhasználásával.

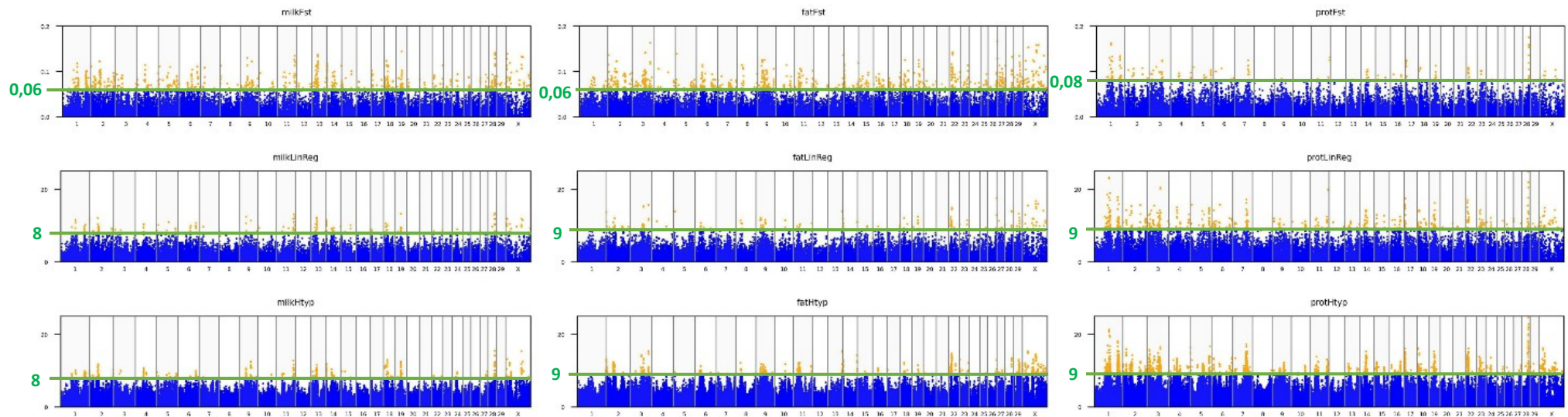
1. ábra. A két vagy három becsült tenyésztértékkel ($TÉ$) kapcsolt egybázisos nukleotidpolimorfizmusok (SNP-k) keresésének folyamatábrája.



Első sor: A Holstein-fríz Tenyésztők Egyesületétől kapott tenyésztértékek TH, ZSH és FH esetében; második sor: az egyes tulajdonságokra alkalmazott tesztek, Fst_mrkr: az SNP-k genetikai távolsága, LinReg: lineáris regresszió, Htype: haplotípus-asszociáció; harmadik sor: a különböző tesztek által azonosított legjobb találatokat (SNP-eket) hozzárendeltük az egyes tulajdonságokhoz. Negyedik sor: az egyidejűleg két vagy három $TÉ$ -kel kapcsolt SNP-eket azonosítottuk.

A haplotípus-asszociációs tesztben a vizsgálatot 5 markerre állítottuk be, és minden egyes haplotípusra χ^2 -tesztet végeztünk. A haplotípusokat az elvárásmaximalizáló (EM, Expectation-Maximization) algoritmus (maximális EM-iteráció = 50, EM-konvergenciátűrés = 0,0001) segítségével konstruáltuk. Az $TÉ_{tej}$, $TÉ_{zsír}$, $TÉ_{fehérje}$ asszociációk Manhattan-diagramjainak vizuális vizsgálata után az Fst_marker és $-\log_{10}(p)$ küszöbértékei a lineáris regresszióban vagy a haplotípus-asszociációban 0,06, 8, illetve 8 voltak a $TÉ_{tej}$ esetében; 0,06, 9, illetve 9 a $TÉ_{zsír}$ esetében; és 0,08, 9, illetve 9 a $TÉ_{fehérje}$ esetében. A küszöbértékek feletti SNP-eket minden egyes tulajdonságra vonatkozóan meghatároztuk, és azonosítottuk azokat, amelyek legalább két tulajdonsággal voltak társíthatók (1. táblázat). Az azonosított 74 SNP téves azonosítási aránya (FDS, False Discovery Rate) $1,3 \times 10^{-21}$ és $6,0 \times 10^{-06}$ között volt. Az Fst_marker, a lineáris regresszió és a haplotípus-asszociációs elemzés eredményeit 0 és 1 értékek közé standardizáltuk, majd az egyes tenyésztértékek esetében átlagoltuk.

2. **ábra.** A tej-, a zsír-, illetve a fehérjetermelésre vonatkozó tenyésztértékek, az Fst_marker, a lineáris regresszió és az öt SNP-haplotípus közötti összefüggések Manhattan-diagramjai



A tej- ($TÉ_{\text{tej}}$; balra), a zsír- ($TÉ_{\text{zsír}}$; közép) és a fehérjetermelésre ($TÉ_{\text{fehérje}}$; jobbra) vonatkozó tenyésztértékek, valamint az Fst_marker (felső sor), a lineáris regresszió (középső sor) és az öt SNP haplotípus (alsó sor) közötti összefüggések Manhattan-diagramjai.

A zöld vonalak azok a küszöbértékek, amelyek felett a markerek toptalálatnak minősülnek (Fst_marker; LinReg; Htyp értékek rendre: $TÉ_{\text{tej}}$: 0,06; 8; 8. $TÉ_{\text{zsír}}$: 0,06; 9; 9. $TÉ_{\text{fehérje}}$: 0,08; 9; 9).

A *Bos taurus* genom ARS-UCD1.2 változatát használtuk a közös találatoktól ± 1 millió bázispár (Mbp) távolságra található gének meghatározására (1. melléklet). Ha egyetlen gént sem találtunk a ± 1 Mbp-n belül, a távolságot kiterjesztettük ± 3 Mbp-ra.

1. táblázat. A legalább két vizsgált paraméterrel összefüggésbe hozható 74 SNP

	SNP-k száma	TÉ _{tej}	TÉ _{zsír}	TÉ _{fehérje}
	5	+	+	
	44	+		+
	16		+	+
	9	+	+	+
<i>Összesen</i>	74	58	30	69

2.2. BLUP-tenyésztéértébecslési módszerek összehasonlítása

A hazai tejágazatban meghatározóak az intenzíven tartott holstein-fríz tehénállományok, ahol egy gazdaságban átlagosan 453 termelésellenőrzött, törzskönyvi nyilvántartásban szereplő tehén termel. Szűrést alkalmaztunk, hogy csak olyan állatok kerüljenek a vizsgálatba, amelyeknél rendelkezésre áll a TÉ_{Pedigri} és a TÉ_{Genomikai}. A saját teljesítményeken alapuló BLUP EM-tenyésztéértébecslést a korábban genotipizált nőivarú egyedek ellése és első laktációjuk zárását, valamint a hivatalos küllemi bírálatukat követően végeztük el. 1 616 549 nőivarú egyed kapott TÉ_{Hagyományos} értéket, közülük korábban 23 561 nőivarú egyed genomikai tenyésztéértékét is meghatároztuk. További szűréseink az azonos telep/év/évszak/életkor paraméterekre vonatkoztak, így végül 190 ugyanabból a gazdaságból származó, azonos életkorú és laktációs stádiumú tehén rendelkezett a teljes adatbázissal és került be az elemzésbe. A genotipizálást követően a TÉ_{Genomikai} becslését a CRV-Holland és a HFTE együttesen végezte el. A csoport hivatalos tejtermelésellenőrzési eredményeinek átlagértéke a következők szerint alakult: 305 napos standard laktációs tejhozam: 10 910 kg (TH, kg), 305 napos zsírhozam: 398 kg (ZSH, kg) és 305 napos fehérjehozam: 365 (FH, kg). A vizsgálatba vont teheneket ugyanabban a gazdaságban, szabad tartásban, pihenőboxokkal ellátott istállóban tartották, és komplett takarmányadag (TMR, Total Mixed Ration) biztosította a táplálóanyag-ellátásukat, amely elsősorban kukoricaszilázsából állt, koncentrált takarmánnyal és egyéb adalékanyagokkal kiegészítve. Ez az azonos tartási és takarmányozási rendszer kulcsfontosságú volt a környezeti különbségek minimalizálásához, továbbá annak biztosításához, hogy a termelési jellemzők és a genetikai értékelések valóban tükrözzék a genetikai különbségeket.

Az összehasonlított háromféle BLUP-módszer az alábbi:

A **pedigré-tenyésztérték** ($TÉ_{Pedigré}$) az anya ($BLUP_{tehén}$) és az apa ($BLUP_{bika}$) hagyományos tenyésztértékének ($TÉ_{Hagyományos}$) egyszerű számtani átlagértékeként számítottuk ki az alábbiak szerint:

$$TÉ_{Pedigré} = \frac{TÉ_{BLUP_{tehén}} + TÉ_{BLUP_{bika}}}{2}$$

A **hagyományos BLUP-tenyésztérték** ($TÉ_{BLUP}$): esetében két mátrix került létrehozásra. Az egyik az adatbázismátrix, a másik pedig a pedigrémátrix. A rokonok pedigrémátrixa tartalmazta a teljes testvérek, féltestvérek, apák, anyák és nagyszülők pedigréadatait. A BLUP-modellek magukban foglalták az anyai genetikai hatásokat és az anyai állandó környezeti hatásokat, mint véletlen hatásokat. A modellek az alábbiak szerint kerültek felépítésre:

$$\mathbf{y} = X_{\mathbf{b}} + Z_{\mathbf{a}} + W_{\mathbf{pe}} + \mathbf{e}$$

ahol \mathbf{y} az észlelések vektora; \mathbf{b} a fix hatások vektora; \mathbf{a} az állatok véletlen hatásainak vektora; \mathbf{pe} az állandó környezeti véletlen hatások vektora; \mathbf{e} a véletlen reziduális hatások vektora; X , Z és W pedig a fix, állati és véletlen állandó környezeti hatásokkal kapcsolódó előfordulási mátrixok.

A **genomikai tenyésztértékbecslés** ($TÉ_{Genomikai}$) módszere a 2.1. pontban került ismertetésre

A termelési tulajdonságok normalitásának értékelésére a Kolmogorov–Smirnov-tesztet alkalmaztuk, míg a változók homogenitásának értékelésére Levene-tesztet használtunk. Az említett tulajdonságokra többszörös varianciaanalízist (ANOVA) végeztünk. Továbbá Pearson-féle és Spearman-féle rangkorreláció segítségével határoztunk meg a termelési mutatók, valamint a különböző tenyésztértékadatok közötti összefüggéseket. Az adatok előkészítésére, rendszerezésére a Microsoft Office szoftvercsomagot (Excel 2019 és Word 2019) használtuk. Az adatbázis kiértékelését az SPSS statisztikai szoftvercsomag 27.0 verziójával végeztük.

3. EREDMÉNYEK ÉS ÉRTÉKELÉSÜK

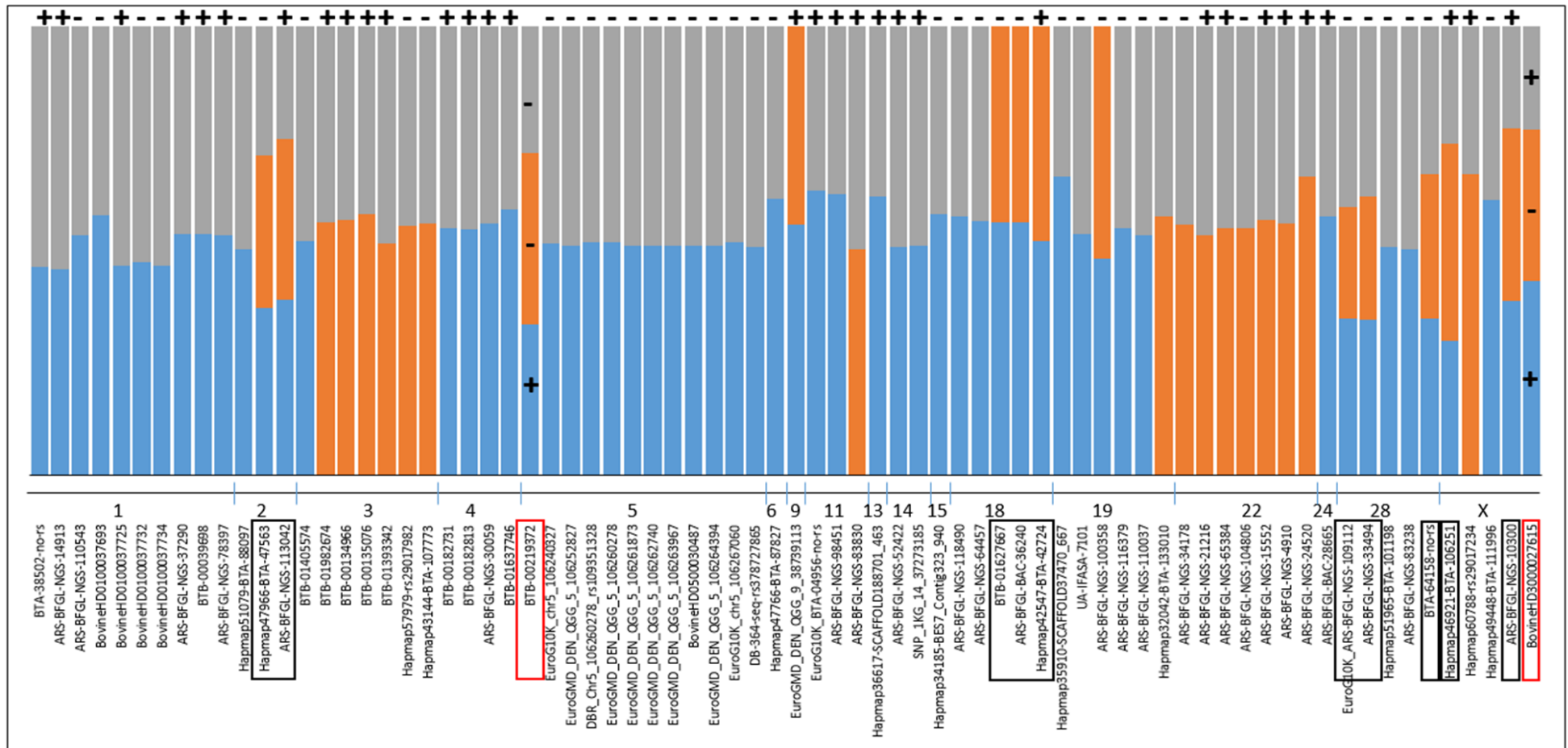
3.1. SNP-vizsgálatok eredménye

A TH és ZSH mutatóval összefüggésbe hozható 5 SNP a 9., 18. és 19. kromoszómán helyezkedett el. Ezenkívül 44 SNP kapcsolódott a TH és a FH tenyésztértékekhez az 1–6, 11, 13–15, 18, 19, 24, 28 és X kromoszómán. Továbbá 16 SNP kapcsolódott a ZSH és a FH tenyésztértékhez a 3, 11, 19, 22 és X BTA-n. Emellett 9 SNP kapcsolódott a TH, ZSH és FH mutatóhoz a 2, 5, 28 és X kromoszómán.

Az azonosított SNP-k maximális értékei a F_{st_marker} esetében 0,17, a lineáris regresszió $-\log_{10}(p)$ értéke 24,9, a haplotípus-asszociáció $-\log_{10}(p)$ értéke 26,4 volt. Eredményeink közül a legkiemelkedőbb találatok – az SNP-k regressziós β -érték átlaga $> 0,8$ – a TH esetében a 2., 11., 19., 28. és X kromoszómán, a ZSH esetében a 3., 22., 28. és X kromoszómán, a FH esetében pedig az 1. és 28. kromoszómán helyezkedtek el.

A TH, ZSH és FH esetében a tenyésztértékekkel összefüggésbe hozott kilenc SNP közül hét a legnagyobb hatásúak között található, három közülük 1,18 millió bázispáron belül helyezkedik el a 28-as kromoszómán (2. melléklet).

3. ábra. A két vagy három tenyésztértekhez kapcsolódó 74 SNP ábrázolása



A narancssárga ($TÉ_{teji}$), kék ($TÉ_{zsi}$) és szürke ($TÉ_{fehérje}$) sávok megfelelnek az egyes SNP-k regressziós β -együtthatóinak (2. melléklet). A mindhárom $TÉ$ -kel összefüggő SNP-eket bekereteztük. Az ábra tetején plusz és mínusz jelöli az összes tulajdonság esetében pozitív vagy negatív β -együtthatót. A piros keret olyan SNP-eket jelöl, ahol a β -együtthatók ellentétesek voltak, például a ZSH és a FH esetében negatív, míg a TH-nál pozitív (BTA 5), vagy a ZSH-nál negatív, a TH és a FH esetében pedig pozitív (X kromoszóma).

3.2. BLUP-tenyésztéértébecslési módszerek összehasonlító eredménye

A vizsgált populáció standard laktációs termelése kedvezően alakult, és megfelelt az országos átlagnak.: 10 910,50 kg tej, 397,86 kg zsír és 365,33 kg fehérje. A 190 tehénből álló csoport homogénnek bizonyult a termelési jellemzők tekintetében, amit a variációs együttható (CV%) 15% alatti értékei is alátámasztottak.

2. táblázat. A vizsgálatba vont állatok statisztikai adatai

Tulajdonság	TH (kg)	ZSH (kg)	FH (kg)
Vizsgált tehenek száma (<i>n</i>)	190	190	190
Közéérték	10 910,50	397,86	365,33
SD	1 453,70	45,24	39,80
CV%	13,32	11,37	10,90
SE	105,46	3,28	2,89
Min.	6 505	260	196
Max.	13 781	511	451

TH = 305 napos tejhozam; ZSH = 305 napos zsírhozam; FH = 305 napos fehérjehozam

A 3. táblázat adatai jelentős különbséget mutatnak a különböző tenyésztéértékek minimális és maximális értékei között; az átlagértékekhez viszonyított standard hibák (SE) azonban viszonylag alacsonyak (5% alatt). A $TÉ_{Pedigré}$ értékei minden esetben magasabbak a másik két $TÉ$ -nél. Ez az eltérés a tenyésztéértébecslésben a szülői generáció és az utódgeneráció között ötévente végrehajtott fokozatos bázisváltásnak tulajdonítható.

3. táblázat. A vizsgált tenyésztéértékek statisztikai adatai

$TÉ$	<i>n</i>	Közéérték	SE	Min.	Max.
TH (kg)					
$TÉ_{Hagyományos}$	190	739,16	31,55	-397	1779
$TÉ_{Genomikai}$	190	718,11	32,31	-357	2052
$TÉ_{Pedigré}$	170	1 379,57	29,98	-186	2354
ZSH (kg)					
$TÉ_{Hagyományos}$	190	34,47	1,00	0	65
$TÉ_{Genomikai}$	190	33,44	1,24	-24	75
$TÉ_{Pedigré}$	170	55,58	0,97	22	87
FH (kg)					
$TÉ_{Hagyományos}$	190	27,01	0,78	-2	56
$TÉ_{Genomikai}$	190	26,05	0,99	-18	66
$TÉ_{Pedigré}$	170	49,33	0,91	6	78

TH = 305 napos tejhozam; ZSH = 305 napos zsírhozam; FH = 305 napos fehérjehozam;
 $TÉ_{Hagyományos}$ = hagyományos BLUP; $TÉ_{Genomikai}$ = genomikai BLUP; $TÉ_{Pedigré}$ = pedigré BLUP

A 4. táblázat összefoglalja a T \acute{E} -ek és a fenotípusos teljesítmény, valamint a három különböző módon becsült T \acute{E} közötti korrelációs együtthatókat. A korrelációs együtthatók értékei minden esetben pozitívak, és többségük szignifikáns. Amint a táblázatból látható, a T \acute{E} _{Hagyományos} szorosabb kapcsolatot mutat a fenotípusos teljesítményekkel ($r_{gp} = 0,61-0,70$), mint a T \acute{E} _{Genomikai} ($r_{gp} = 0,31-0,48$). A leggyengébb összefüggést ($r_{gp} = 0,15-0,24$) a genotípus és a fenotípus között a T \acute{E} _{Pedigré} esetében találtuk. A T \acute{E} _{Genomikai} mérsékelt és szoros asszociációt mutatott ($r_g = 0,66-0,67$) a T \acute{E} _{Hagyományos}-sal. A T \acute{E} _{Hagyományos} és a T \acute{E} _{Pedigré} között szintén közepes és szoros a kapcsolat ($r_g = 0,56-0,66$), a T \acute{E} _{Genomikai} és a T \acute{E} _{Pedigré} között pedig valamivel gyengébb, mérsékelt vagy szoros ($r_g = 0,43-0,56$).

4. táblázat. A fenotípusos teljesítmény, a tenyésztékek és a három különböző módon becsült T \acute{E} közötti korrelációs együtthatók

Korreláció (r)	T \acute{E} _{Hagyományos}	T \acute{E} _{Genomikai}	T \acute{E} _{Pedigré}
TEJ	$r_{gp} = 0,70; p < 0,01$	$r_{gp} = 0,48; p < 0,01$	$r_{gp} = 0,24; p < 0,01$
T \acute{E} _{Hagyományos}		$r_g = 0,67; p < 0,01$ $r_{rank} = 0,65; p < 0,01$	$r_g = 0,66; p < 0,01$ $r_{rank} = 0,50; p < 0,01$
T \acute{E} _{Genomikai}			$r_g = 0,53; p < 0,01$ $r_{rank} = 0,40; p < 0,01$
ZSÍR	$r_{gp} = 0,69; p < 0,01$	$r_{gp} = 0,32; p < 0,01$	$r_{gp} = 0,15; NS$
T \acute{E} _{Hagyományos}		$r_g = 0,67; p < 0,01$ $r_{rank} = 0,65; p < 0,01$	$r_g = 0,56; p < 0,01$ $r_{rank} = 0,57; p < 0,01$
T \acute{E} _{Genomikai}			$r_g = 0,43; p < 0,01$ $r_{rank} = 0,41; p < 0,01$
FEHÉRJE	$r_{gp} = 0,61; p < 0,01$	$r_{gp} = 0,31; p < 0,01$	$r_{gp} = 0,15; p < 0,05$
T \acute{E} _{Hagyományos}		$r_g = 0,66; p < 0,01$ $r_{rank} = 0,66; p < 0,01$	$r_g = 0,60; p < 0,01$ $r_{rank} = 0,12; NS$
T \acute{E} _{Genomikai}			$r_g = 0,56; p < 0,01$ $r_{rank} = 0,56; p < 0,01$

T \acute{E} _{Hagyományos} = hagyományos BLUP EM; T \acute{E} _{Genomikai} = genomikai BLUP;

T \acute{E} _{Pedigré} = pedigré BLUP; r_{gp} = fenotípus-genotípus korreláció;

r_g = genetikai korreláció; r_{rank} = rangkorreláció

A T \acute{E} _{Hagyományos} és a T \acute{E} _{Genomikai} közötti rangkorrelációs értékek ($r_{rank} = 0,65-0,66$) szoros kapcsolatot jeleznek. A T \acute{E} _{Hagyományos} és a T \acute{E} _{Pedigré} közötti korrelációs együtthatók ($r_{rank} = 0,12-0,57$), továbbá a T \acute{E} _{Genomikai} és a T \acute{E} _{Pedigré} közötti értékek is közepes szorosságú összefüggést mutatnak ($r_{rank} = 0,40-0,56$).

A T \acute{E} -ek fenotípusos tulajdonságok összefüggésére irányuló regressziós elemzéseinek eredményeit az 5. táblázat mutatja be. A ZSH esetében a T \acute{E} _{Pedig \acute{r} e} kivételével szignifikáns ($p < 0,01$) összefüggéseket találtunk. Valamennyi regressziós együttható (b) értéke pozitív, a legmagasabb a T \acute{E} _{Hagyományos} esetében ($b = 0,17-0,21$), majd a T \acute{E} _{Genomikai} ($b = 0,10-0,15$) és legalacsonyabb értékkel a T \acute{E} _{Pedig \acute{r} e} ($b = 0,04-0,07$) következik. Az adatok egyértelműen azt mutatják, hogy a legnagyobb pontosságot ($R^2 = 0,37-0,48$) minden fenotípusos tulajdonságra a T \acute{E} _{Hagyományos} éri el, ezt követi a T \acute{E} _{Genomikai} ($R^2 = 0,09-0,23$), a legkevésbé pontosnak pedig a T \acute{E} _{Pedig \acute{r} e} bizonyul ($R^2 = 0,02-0,06$). A termelési tulajdonságok közül a TH-ra és a ZSH-ra a hagyományos T \acute{E} -ek nagyobb hatást gyakorolnak, mint a FH-ra. A hatás ($b = 0,17-0,21$) minden esetben a pozitív irányba mutat, a T \acute{E} _{Hagyományos} esetében a legmagasabb, ezt sorrendben a T \acute{E} _{Genomikai} (0,10–0,15) és a T \acute{E} _{Pedig \acute{r} e} követi (0,04–0,07).

5. táblázat. A regressziós vizsgálat eredményei

T \acute{E} (Y)	Tulajdonság (X)	Merekség			Tengelymetszet			Illeszkedés	
		b	SE	p	a	SE	p	R^2	p
T \acute{E} _{Hagyományos}	TH	0,21	0,02	< 0,01	-1 529,56	172,66	< 0,01	0,48	< 0,01
T \acute{E} _{Genomikai}	TH	0,15	0,02	< 0,01	-879,32	216,02	< 0,01	0,23	< 0,01
T \acute{E} _{Pedig\acute{r}e}	TH	0,07	0,02	< 0,01	665,15	229,32	< 0,01	0,06	< 0,01
T \acute{E} _{Hagyományos}	ZSH	0,21	0,02	< 0,01	-48,80	6,51	< 0,01	0,47	< 0,01
T \acute{E} _{Genomikai}	ZSH	0,12	0,03	< 0,01	-14,20	10,47	< 0,01	0,10	< 0,01
T \acute{E} _{Pedig\acute{r}e}	ZSH	0,04	0,02	NS	38,83	8,66	< 0,01	0,02	NS
T \acute{E} _{Hagyományos}	FH	0,17	0,02	< 0,01	-33,12	5,75	< 0,01	0,37	< 0,01
T \acute{E} _{Genomikai}	FH	0,10	0,02	< 0,01	-12,12	8,73	< 0,01	0,09	< 0,01
T \acute{E} _{Pedig\acute{r}e}	FH	0,05	0,02	< 0,05	32,06	8,77	< 0,01	0,02	< 0,05

TH = 305 napos tejhozam; ZSH = 305 napos zsírhozam; FH = 305 napos fehérjehozam;
T \acute{E} _{Hagyományos} = hagyományos BLUP EM; T \acute{E} _{Genomikai} = genomikai BLUP; T \acute{E} _{Pedig \acute{r} e} = pedig \acute{r} e BLUP

Eredményeink- jelentős különbségeket tártak fel a becsült tenyésztési értékek pontosságában és megbízhatóságában azonos évben született, azonos körülmények között, ugyanabban a gazdaságban és állományban felnevelt, első laktációs tejelő tehenek esetében Magyarországon. Ezek a megállapítások rávilágítanak az összehasonlító elemzések fontosságára. Az ilyen vizsgálatok elengedhetetlenek a tenyésztési értékebecslések pontosságának növeléséhez, különösen ellenőrzött környezeti feltételek mellett.

Az a megállapítás, hogy a T \acute{E} _{Hagyományos}-módszer mutatja a legnagyobb megbízhatóságot, amelyet a T \acute{E} _{Genomikai} és végül a T \acute{E} _{Pedig \acute{r} e} követ, nem csökkenti a széles körben elfogadott T \acute{E} _{Genomikai}-becslés jelentőségét. Vizsgálatunk eredményei rámutatnak a különböző becslési módszerek eltérő megbízhatóságára azonos környezeti feltételek mellett. Ez a sorrend kiemeli a genetikai értékelések optimalizálásának és a tenyésztési stratégiák pontosságának növelésére irányuló lehetőségeket, ami a hatékonyabb tenyésztési programok alapját képezheti.

A $TÉ_{\text{Genomikai}}$ -módszer előnye a $TÉ_{\text{Hagyományos}}$ -hoz képest kevésbé volt nyilvánvaló kis populációméretben vagy azonos környezeti feltételek mellett. Ugyanakkor a $TÉ_{\text{Genomikai}}$ -módszer jelentős előnyt kínál olyan nehezen becsülhető tulajdonságok esetében, mint például a menedzsmenttel összefüggő jellemzők, a hosszú hasznos élettartam, az életteljesítmény, valamint különböző élettani (metabolikus és egészségügyi) tulajdonságok.

A genomikai becslés korai életszakaszban történő alkalmazása lehetővé teszi a tenyésztők számára, hogy értékes időt takarítsanak meg, és csökkentsék a felesleges fenntartási költségeket az állomány méretének optimalizálásával. A $TÉ_{\text{Genomikai}}$ -eredményekre alapozott folyamatos tenyésztésintű szelekció jelentősen javította az állomány genetikai minőségét, elősegítve a kedvező genetikai tulajdonságokkal rendelkező egyedek kiválasztását és továbbtenyésztését. A $TÉ_{\text{Hagyományos}}$ -módszer pedig a $TÉ_{\text{Genomikai}}$ kontrolljaként szolgálhat állományi szinten, tovább növelve ezzel a genetikai értékelések megbízhatóságát.

4. ÚJ TUDOMÁNYOS EREDMÉNYEK ÉS HASZNOSÍTÁSUK

4.1. Új tudományos eredmények

1. Elsőként alkalmaztam a fenotípus és a genotípus közötti Pearson-féle korrelációt a különböző tenyésztértékbecslési módszerek megbízhatóságának jellemzésére. A három BLUP-módszer összehasonlítása alapján megállapítottam, hogy a hagyományos BLUP nyújtotta a legmegbízhatóbb becsléseket ($r = 0,61-0,70$), ezt követte a genomikai adatokkal kiegészített BLUP ($r = 0,31-0,48$), míg a legkevésbé megbízhatónak a pedigré BLUP bizonyult ($r = 0,15-0,24$).
2. Ráműtöttem arra, hogy a holstein-fríz tehének termelési tulajdonságokra becsült tenyésztértékei validálására alkalmas az azonos évben, évszakban született, azonos környezetben tartott, azonos takarmányozásban, gondozásban részesített egyedek 305 napos tej-, tejszír- és tejfehérje-termelése.
3. Elsőként határoztam meg holstein-fríz fajtában három különböző módszer (az SNP-k genetikai távolsága, lineáris regresszió és haplotípus-asszociáció) alapján olyan SNP-eket, amelyek egyidejűleg két vagy három – tejtermeléssel kapcsolatos – tenyésztérték kialakításában is szerepet játszanak.
4. Az azonosított SNP-k (74) regressziós vizsgálata alapján megállapítottam, hogy a legtöbb marker tenyésztértékekre gyakorolt hatása azonos irányba mutatott. Ugyanakkor az elemzés során olyan markereket is azonosítottam (BTB-00219372 és BovineHD3000027615), amelyek alkalmazása óvatosságot igényel, mivel ezek míg az egyik tenyésztértéket növelik, addig egy vagy akár mindkét másik tenyésztérték csökkenését eredményezhetik.
5. Négy olyan gént azonosítottam (*EFCAB10*, *GLOD5*, *NONO* és *TMEM70*), amelyeket korábban nem vizsgáltak szarvasmarhában, illetve nem hozták összefüggésbe tejtermelési tulajdonságokkal.

4.2. Az eredmények hasznosíthatósága

A felelős holstein-fríz tenyésztés új megközelítést kíván, amely számos tényező, köztük a tenyészértékek együttes figyelembevételét helyezi előtérbe a fenntarthatóság és a regeneratív fejlődés érdekében. A genotípus és a környezet közötti egyensúly fenntartása jelentős mértékben javíthatja a HF fajtára alapozott tejtermelés fenntarthatóságát és alkalmazkodóképességét. A genomikai információ és a genomikai szelekciós rendszerek alkalmazása lehetőséget teremt arra, hogy az ezekre alapozott tudatos tenyésztési döntések eredményeként a tehének egészségesebben, jobb szaporodásbiológia mutatók mellett, hatékonyabban, kevesebb takarmányból több tejet termeljenek, ezzel is csökkentve a környezetre gyakorolt kedvezőtlen hatásokat.

A vizsgálat során azonosított SNP-k számos olyan gén közelében helyezkednek el, amelyeket szarvasmarhák esetében eddig nem vizsgáltak, és amelyek potenciális célpontként szolgálhatnak a tejtermelési mutatókhoz kapcsolódó újabb gyakorlati alkalmazásokhoz. A két vagy három tenyészértékkel összefüggő markerek hatékonyan alkalmazhatók a szelekcióban, elősegítve a genetikai fejlődés felgyorsítását, valamint támogatva a tenyésztőket, szakembereket, szarvasmarhatartókat tenyésztési és szelekciós céljaik elérésében. Az azonos irányba ható markerek megkönnyítik a tenyész kiválasztást, míg az ellentétes hatású markerek kezelése fokozott körültekintést igényel, mivel ezek egy tenyészértéket növelhetnek, miközben egy másikat vagy akár kettőt is csökkenthetnek.

A vizsgálati eredmények értékét a tenyészértékek validálásában való alkalmazhatóságuk adja, különösen az eltérő becslési módszerek és a ténylegesen megvalósult teljesítmény közvetlen összehasonlítása révén, egységes környezeti feltételek mellett. Ez új perspektívát kínál a genomikai szelekció gyakorlati alkalmazásához a HF szarvasmarha tenyésztésében.

A kutatás eredményei kézzelfogható gazdasági előnyt nyújtanak a hazai HF tenyésztők számára, mivel a korai tenyész kiválasztás és a hatékony (nemzeti, illetve telepi szintű) tenyésztési programok hozzájárulhatnak az optimális állomány méret kialakításához, ezáltal növelve a tejtermelés jövedelmezőségét, és jelentősen csökkentve a környezet terhelését.

5. AZ ÉRTEKEZÉS TÉMAKÖRÉBEN MEGJELENT PUBLIKÁCIÓK

5.1. Idegen nyelvű folyóiratban megjelent lektorált cikk

Bognár, L. – Kőrösi, Zs. J. – Anton, I. – Bene, Sz. - Szabó, F. (2024): Different Breeding Values Under Uniform Environmental Condition for Milk Production Yield Traits in Holstein-Friesian Cows. *Animals* 15 (51), 11. <https://doi.org/10.3390/ani15010051> Q1, IF=2,7

Bognár, L. – Kőrösi, Zs. J. – Bene, Sz. A. – Szabó, F. – Anton, I. - Zsolnai, A. (2024): Simultaneous Effects of Single-Nucleotide Polymorphisms on the Estimated Breeding Value for Milk, Fat, and Protein Yield of Holstein-Friesian Cows in Hungary. *Animals* 14 (23), 3518. <https://doi.org/10.3390/ani14233518> Q1, IF=2,7

Bene, Sz. – Kőrösi, Zs. J. – **Bognár, L.** – Polgár, J. P. – Szabó, F. (2024): Population Genetic Features of Calving Interval of Holstein-Friesian Cows Bred in Hungary. *Animals* 14 (17), 2513. <https://doi.org/10.3390/ani14172513> Q1, IF=2,7

Bognár, L. – Szabó, F. (2023): Management of “Modern” Holstein Cows Focusing on Sustainability and Resilience – Review of Recent Achievements. *Chemical Engineering Transactions* 107, pp. 169–174. Q3

5.2. Magyar nyelvű folyóiratban megjelent lektorált cikk

Bognár L. - Kőrösi Zs. J. - Bene Sz. - Szabó F. (2024): A modern holstein-fríz tehén tenyésztésének és tartásának néhány fenntarthatósági szempontja (Irodalmi áttekintés) (Some sustainability aspects for the breeding and management of modern Holstein cow. [Literature review]) *Állattenyésztés és Takarmányozás* 73 (2), 109–119. old.

Bognár L. (2021): A hazai genomikai tenyészértékbecslés gyakorlati tapasztalatai a holstein-fríz fajta tenyésztési programjában (Practical Experiences of the Hungarian Genomic Breeding Value Estimation in the Breeding Program of the Holstein-Friesian Genetic Programme). *Állattenyésztés és Takarmányozás*, 70 (3), 250–269. old.

Bognár L. (2013): Új irányzatok a tejtermelő szarvasmarha-tenyésztésben. A genomikus tenyészértékbecslés. (New Trends in Dairy Cattle Breeding. Genomic Breeding Value Estimation). *Állattenyésztés és Takarmányozás*, 62 (4), 367–373. old.

6. MELLÉKLETEK

6.1. Melléklet – *B. taurus* genom ARS-UCD1.2

A két vagy három TÉ-hez kapcsolódó markerek neve, genomialis pozíciójuk és a markerek körül található gének.

Marker	Kromoszóma	Pozíció (bp)	Tej	TÉ		± 1 millió bázispár	± 3 millió bázispár
				Zsír	Fehérje		
BTA-38502-no-rs	1	82 339 579	+		+	<i>CIH3orf70</i>	
ARS-BFGL-NGS-14913	1	82 360 713	+		+	<i>LINC02054,</i> <i>CLCN2,</i> <i>MAP6D1,</i> <i>YEATS2</i>	
ARS-BFGL-NGS-110543	1	86 233 203	+		+	<i>TTC14tm</i>	
BovineHD0100037693	1	131 261 458	+		+	<i>ARMC8, CEP70</i>	
BovineHD0100037725	1	131 389 856	+		+		
BovineHD0100037732	1	131 413 283	+		+		
BovineHD0100037734	1	131 427 641	+		+		
ARS-BFGL-NGS-37290	1	136 049 399	+		+	<i>TOPBP1</i>	
BTB-00039698	1	136 090 182	+		+		
ARS-BFGL-NGS-78397	1	137 305 479	+		+		
Hapmap51079-BTA-88097	2	20 011 118	+		+	<i>HOXD3</i>	
Hapmap47966-BTA-47563	2	48 824 025	+	+	+		
ARS-BFGL-NGS-113042	2	48 993 143	+	+	+		
BTB-01405574	3	40 190 639	+		+	<i>RNPC3</i>	
BTB-01982674	3	61 970 189		+	+		
BTB-00134966	3	69 934 639		+	+		
BTB-00135076	3	69 961 802		+	+		
BTB-01393342	3	71 022 745		+	+		
Hapmap57979-rs29017982	3	73 879 769		+	+		
Hapmap43144-BTA-107773	3	73 904 881		+	+	<i>CTH</i>	
BTB-00182731	4	46 190 578	+		+		<i>PHTF2</i>
BTB-00182813	4	46 311 240	+		+	<u><i>EFCAB10</i></u>	
ARS-BFGL-NGS-30059	4	46 943 050	+		+		
BTB-01637746	4	47 751 036	+		+		

BTB-00219372	5	9 085 964	+	+	+	<i>PPFIA2,</i> <i>METTL25</i>
EuroG10K_chr5_106240327	5	105 749 785	+		+	<i>TAPBPL*</i>
EuroGMD_DEN_QGG_5_106252827	5	105 762 284	+		+	
DBR_Chr5_106260278_rs109351328	5	105 769 735	+		+	
EuroGMD_DEN_QGG_5_106260278	5	105 769 735	+		+	
EuroGMD_DEN_QGG_5_106261873	5	105 771 330	+		+	
EuroGMD_DEN_QGG_5_106262740	5	105 772 197	+		+	
EuroGMD_DEN_QGG_5_106263967	5	105 773 382	+		+	
BovineHD0500030487	5	105 773 809	+		+	
EuroGMD_DEN_QGG_5_106264394	5	105 773 809	+		+	
EuroG10K_chr5_106267060	5	105 776 475	+		+	
DB-364-seq-rs378727865	5	105 784 987	+		+	<i>TSPAN11</i> <i>FBXL14</i>
Hapmap47766-BTA-87827	6	100 139 940	+		+	<i>HNRNPD,</i> <i>HNRNPDL,</i> <i>KLHL8</i>
EuroGMD_DEN_QGG_9_38739113	9	38 271 438	+	+		<i>TUBE1,</i> <i>MFSD4B</i>
EuroG10K_BTA-04956-no-rs	11	94 715 801	+		+	<i>STRBP</i>
ARS-BFGL-NGS-98451	11	97 225 356	+		+	<i>STRBP</i>
ARS-BFGL-NGS-83830	11	102 752 125		+	+	<i>NTNG2, TTF1,</i> <i>SPACA9^{cfm},</i> <i>PAEP, KCNT1,</i> <i>CARD9</i>
Hapmap36617- SCAFFOLD188701_463	13	34 319 303	+		+	<i>PTCHD3</i>
ARS-BFGL-NGS-52422	14	37 166 868	+		+	<i>SBSPON,</i> <i>TMEM70, JPH1</i>
SNP_1KG_14_37273185	14	37 273 185	+		+	
Hapmap34185-BES7_Contig323_940	15	55 356 241	+		+	<i>ARAP1, P2RY2,</i> <i>P4HA3*, MYO7A</i>
ARS-BFGL-NGS-118490	18	12 718 603	+		+	<i>EMC8</i>
ARS-BFGL-NGS-64457	18	12 739 390	+		+	
BTB-01627667	18	21 596 766	+	+		<i>NOD2</i>
ARS-BFGL-BAC-36240	18	21 636 399	+	+		
Hapmap42547-BTA-42724	18	21 791 294	+	+		

Hapmap35910- SCAFFOLD37470_667	19	31 569 771	+	+		
UA-IFASA-7101	19	32 389 986	+	+		
ARS-BFGL-NGS-100358	19	32 754 596	+	+		
ARS-BFGL-NGS-116379	19	32 997 278	+	+		
ARS-BFGL-NGS-110037	19	33 052 413	+	+	<i>ZNF624,</i> <i>CENPV, ULK2</i>	
Hapmap32042-BTA-133010	19	33 716 989		+	+	<i>ALDH3A2,</i> <i>B9D1^{efm},</i> <i>PRPSAP2,</i> <i>MYO15A,</i> <i>TOM1L2</i>
ARS-BFGL-NGS-34178	22	10 597 656		+	+	<i>C22H3orf35,</i> <i>DLEC1^{efm}</i>
ARS-BFGL-NGS-21216	22	11 596 830		+	+	<i>MYD88</i>
ARS-BFGL-NGS-65384	22	11 862 971		+	+	<i>SLC22A13,</i> <i>ACVR2B, EXOG</i>
ARS-BFGL-NGS-104806	22	12 291 232		+	+	<i>MOBP</i>
ARS-BFGL-NGS-15552	22	14 961 300		+	+	<i>CTNNB1,</i> <i>CCDC13^{efm},</i> <i>ACKR2</i>
ARS-BFGL-NGS-4910	22	16 130 993		+	+	<i>ZNF852,</i> <i>TATDN2,</i> <i>RPUSD3,</i> <i>TTLL3, SETD5</i>
ARS-BFGL-NGS-24520	22	18 761 349		+	+	<i>SRGAP3, OXTR</i>
ARS-BFGL-BAC-28665	24	28 487 771	+	+		<i>ZNF503</i>
EuroG10K_ARS-BFGL-NGS-109112	28	36 015 224	+	+	+	
ARS-BFGL-NGS-33494	28	36 097 359	+	+	+	
Hapmap51965-BTA-101198	28	36 110 502	+	+	+	
ARS-BFGL-NGS-83238	28	36 205 983	+	+	+	
BTA-64158-no-rs	28	37 195 142	+	+	+	<i>CCSER2, SHLD2,</i> <i>ANXA8L1</i>
Hapmap46921-BTA-106251	X	30 978 737	+	+	+	<i>FMR1,</i> <i>FMR1NB, AFF2,</i> <i>IDS,</i> <i>CXHXorf40A,</i> <i>TMEM185A</i>
Hapmap60788-rs29017234	X	77 312 570		+	+	<i>NEXMIF, RLIM, CITED1, RPS4X,</i> <i>SLC16A2, ERCC6L, PIN4,</i>

						<i>ZCCHC13,</i>	<i>NHSL2,</i>	<i>RTL5,</i>
						<i>CHIC1, CDX4,</i>	<i>CXCR3,</i>	<i>GCNA,</i>
						<i>MGC140080,</i>	<i>OGT,</i>	<i>TAFI,</i>
						<i>PABPC1L2A,</i>	<i>ITGB1BP2,</i>	
						<i>PHKA1, HDAC8</i>	<i>NONO,</i>	<i>ZMYM3,</i>
							<i>GJB1,</i>	<i>NLGN3,</i>
							<i>MED12</i>	
Hapmap49448-BTA-111996	X	87 848 657	+	+	+	<i>GLOD5, GATA1,</i>	<i>SLC38A5, FTSJ1,</i>	
						<i>HDAC6, ERAS,</i>	<i>PORCN, EBF,</i>	
						<i>PCSK1N,</i>	<i>TBC1D25, RBM3,</i>	
						<i>TIMM17B,</i>	<i>WDR13, WAS,</i>	
						<i>PQBPI,</i>	<i>SUV39H1</i>	
						<i>SLC35A2, PIM2,</i>		
						<i>OTUD5,</i>		
						<i>KCND1,</i>		
						<i>GRIPAP1, TFE3,</i>		
						<i>CCDC120,</i>		
						<i>PRAF2, WDR45,</i>		
						<i>GPKOW, PLP2,</i>		
						<i>PRICKLE3, SYR,</i>		
						<i>CACNA1F,</i>		
						<i>CCDC22,</i>		
						<i>FOXP3,</i>		
						<i>PPP1R3F,</i>		
						<i>PAGE4,</i>		
						<i>USP27X,</i>		
						<i>CLCN5</i>		
ARS-BFGL-NGS-10300	X	87 915 822	+	+	+	<i>AKAP4, CCNB3,</i>	<i>CXHXorf67,</i>	
						<i>DGKK,</i>	<i>NUDT11, GSPT2,</i>	
						<i>SHROOM4,</i>	<i>MAGED1,</i>	
						<i>BMP15,</i>	<i>MAGED4B</i>	
						<i>NUDT10</i>		
BovineHD3000027615	X	95 636 192	+	+	+	<i>HEPH, VSIG4,</i>	<i>SPIN4, ZXDB,</i>	
						<i>MSN, LASIL,</i>	<i>ZNF674</i>	
						<i>ZC3H12B,</i>		
						<i>ZC4H2, ASB12,</i>		
						<i>AMER1,</i>		
						<i>ARHGEF9</i>		

Megjegyzés: Az egyes SNP-ket vagy egymáshoz közel elhelyezkedő SNP-csoportokat vízszintes vonalak választják el egymástól. Az aláhúzott gének a kalciumtranszportban játszanak szerepet, a csillaggal jelöltek pedig a kollagénnel kapcsolatos folyamatokban. A *cfm* jelölés azokra a génekre vonatkozik, amelyek ismert csilló- és flagellafunkciókkal rendelkeznek és/vagy membránfehérjékként ismertek.

6.2. Melléklet – Az egyes markerekhez kapcsolódó vizsgálati értékek

A táblázat az egyes SNP-k genomiai pozícióit, az alkalmazott algoritmusokkal kapott eredmények 0 és 1 közé történő skálázásával standardizált adatok TH, ZSH és FH tenyésztékekhez tartozó átlagait, valamint az egyes tenyésztékekre illesztett lineáris regressziók β -együtthatóit mutatja

Marker	Kromoszóma	Pozíció (bp)	Átlag			β -érték		
			Tej	Zsír	Fehérje	Tej	Zsír	Fehérje
BTA-38502-no-rs	1	82 339 579	0,7668		0,8850	0,2926		0,3150
ARS-BFGL-NGS-14913	1	82 360 713	0,7427		0,8804	0,2724		0,3006
ARS-BFGL-NGS-110543	1	86 233 203	0,5700		0,4984	-0,2157		-0,2167
BovineHD0100037693	1	131 261 458	0,6135		0,4454	0,2750		0,2249
BovineHD0100037725	1	131 389 856	0,5922		0,6775	-0,2426		-0,3005
BovineHD0100037732	1	131 413 283	0,6472		0,7173	-0,2548		-0,3103
BovineHD0100037734	1	131 427 641	0,6299		0,7224	-0,2478		-0,3087
ARS-BFGL-NGS-37290	1	136 049 399	0,6173		0,5302	0,2585		0,2305
BTB-00039698	1	136 090 182	0,6173		0,5333	0,2585		0,2316
ARS-BFGL-NGS-78397	1	137 305 479	0,6890		0,6020	0,3298		0,2981
Hapmap51079-BTA-88097	2	20 011 118	0,5978		0,5888	-0,2179		-0,2352
Hapmap47966-BTA-47563	2	48 824 025	0,7321	0,6739	0,5653	-0,2508	-0,2331	-0,2368
ARS-BFGL-NGS-113042	2	48 993 143	0,8243	0,7631	0,5296	0,2656	0,2505	0,2254
BTB-01405574	3	40 190 639	0,6453		0,5934	-0,2412		-0,2611
BTB-01982674	3	61 970 189		0,6965	0,5421		0,2209	0,1972
BTB-00134966	3	69 934 639		0,6459	0,4890		0,2328	0,2184
BTB-00135076	3	69 961 802		0,6733	0,4820		0,2381	0,2197
BTB-01393342	3	71 022 745		0,5919	0,5536		0,2557	0,2578
Hapmap57979-rs29017982	3	73 879 769		0,9021	0,7212		-0,2956	-0,2887
Hapmap43144-BTA-107773	3	73 904 881		0,9100	0,7127		-0,2940	-0,2843
BTB-00182731	4	46 190 578	0,6399		0,5215	0,2512		0,2499
BTB-00182813	4	46 311 240	0,6291		0,5190	0,2499		0,2499
ARS-BFGL-NGS-30059	4	46 943 050	0,7054		0,5500	0,2577		0,2545
BTB-01637746	4	47 751 036	0,7527		0,5175	0,2384		0,2278
BTB-00219372	5	9 085 964	0,6956	0,7898	0,5834	0,2332	-0,2725	-0,2362
EuroG10K_chr5_106240327	5	105 749 785	0,5999		0,5597	-0,4027		-0,3961
EuroGMD_DEN_QGG_5_106252827	5	105 762 284	0,5701		0,5477	-0,3634		-0,3856
DBR_Chr5_106260278_rs109351328	5	105 769 735	0,6168		0,5725	-0,4062		-0,4034
EuroGMD_DEN_QGG_5_106260278	5	105 769 735	0,6168		0,5725	-0,4062		-0,4034
EuroGMD_DEN_QGG_5_106261873	5	105 771 330	0,5701		0,5477	-0,3634		-0,3856
EuroGMD_DEN_QGG_5_106262740	5	105 772 197	0,5701		0,5477	-0,3634		-0,3856
EuroGMD_DEN_QGG_5_106263967	5	105 773 382	0,5701		0,5481	-0,3634		-0,3856
BovineHD0500030487	5	105 773 809	0,5701		0,5481	-0,3634		-0,3856
EuroGMD_DEN_QGG_5_106264394	5	105 773 809	0,5701		0,5481	-0,3634		-0,3856
EuroG10K_chr5_106267060	5	105 776 475	0,6168		0,5729	-0,4062		-0,4034
DB-364-seq-rs378727865	5	105 784 987	0,5658		0,5461	-0,3618		-0,3849
Hapmap47766-BTA-87827	6	100 139 940	0,7370		0,4623	-0,2490		-0,2081
EuroGMD_DEN_QGG_9_38739113	9	38 271 438	0,7574	0,6010		0,3173	0,2971	

EuroG10K_BTA-04956-no-rs	11	94 715 801	0,8646		0,5003	0,2578		0,2107
ARS-BFGL-NGS-98451	11	97 225 356	0,8027		0,4776	0,2630		0,2053
ARS-BFGL-NGS-83830	11	102 752 125		0,6571	0,6506		0,2543	0,2778
Hapmap36617- SCAFFOLD188701_463	13	34 319 303	0,7220		0,4398	0,2803		0,2454
ARS-BFGL-NGS-52422	14	37 166 868	0,6053		0,5841	0,2241		0,2257
SNP_1KG_14_37273185	14	37 273 185	0,5976		0,5708	0,2220		0,2254
Hapmap34185-BES7_Contig323_940	15	55 356 241	0,6617		0,4753	-0,3369		-0,3017
ARS-BFGL-NGS-118490	18	12 718 603	0,7949		0,5844	-0,2317		-0,2198
ARS-BFGL-NGS-64457	18	12 739 390	0,7497		0,5761	-0,2215		-0,2169
BTB-01627667	18	21 596 766	0,7656	0,5947		-0,2287	-0,2232	
ARS-BFGL-BAC-36240	18	21 636 399	0,7722	0,5953		-0,2287	-0,2232	
Hapmap42547-BTA-42724	18	21 791 294	0,7931	0,7306		0,2329	0,2502	
Hapmap35910- SCAFFOLD37470_667	19	31 569 771	0,9325		0,4704	-0,2736		-0,2033
UA-IFASA-7101	19	32 389 986	0,5894		0,5096	-0,2218		-0,2185
ARS-BFGL-NGS-100358	19	32 754 596	0,7236	0,7733		-0,2312	-0,2633	
ARS-BFGL-NGS-116379	19	32 997 278	0,6596		0,5390	-0,2281		-0,2378
ARS-BFGL-NGS-110037	19	33 052 413	0,6940		0,6037	-0,2337		-0,2501
Hapmap32042-BTA-133010	19	33 716 989		0,6163	0,4525		-0,3042	-0,2666
ARS-BFGL-NGS-34178	22	10 597 656		0,7130	0,5641		-0,2602	-0,2403
ARS-BFGL-NGS-21216	22	11 596 830		0,6881	0,5993		0,2580	0,2556
ARS-BFGL-NGS-65384	22	11 862 971		0,7445	0,6077		0,2656	-0,2572
ARS-BFGL-NGS-104806	22	12 291 232		0,6404	0,5233		-0,2363	-0,2280
ARS-BFGL-NGS-15552	22	14 961 300		0,6521	0,4941		0,2798	0,2369
ARS-BFGL-NGS-4910	22	16 130 993		0,6651	0,5185		0,2643	0,2294
ARS-BFGL-NGS-24520	22	18 761 349		0,8991	0,4542		0,3031	0,2328
ARS-BFGL-BAC-28665	24	28 487 771	0,6145		0,4514	0,2651		0,2479
EuroG10K_ARS-BFGL-NGS-109112	28	36 015 224	0,8100	0,5828	0,9359	-0,2848	-0,2371	-0,3302
ARS-BFGL-NGS-33494	28	36 097 359	0,8524	0,6801	0,9310	-0,2561	-0,2410	-0,2988
Hapmap51965-BTA-101198	28	36 110 502	0,8563		0,8292	-0,2521		-0,2616
ARS-BFGL-NGS-83238	28	36 205 983	0,6997		0,6887	-0,2498		-0,2493
BTA-64158-no-rs	28	37 195 142	0,9260	0,8600	0,8761	-0,2722	-0,2762	-0,2726
Hapmap46921-BTA-106251	X	30 978 737	0,6630	0,9791	0,5772	0,2263	0,2857	0,2338
Hapmap60788-rs29017234	X	77 312 570		0,9058	0,4445		0,3222	0,2439
Hapmap49448-BTA-111996	X	87 848 657	0,7747		0,4882	-0,2287		-0,2012
ARS-BFGL-NGS-10300	X	87 915 822	0,9348	0,9237	0,5436	0,2584	0,2952	0,2282
BovineHD3000027615	X	95 636 192	0,8607	0,6759	0,4581	0,2432	-0,2187	0,1896

Megjegyzés: Az alkalmazott algoritmusokkal kapott eredmények standardizált értékeinek átlagai közül a 0,8-nél nagyobbakat félkövér betűvel jeleztük. Az olyan SNP-k regressziós β -értékei, amelyek hatásiránya különbözött a TH, ZSH és FH tenyésztékek tekintetében, szintén félkövér kiemélést kaptak.

6.3. Melléklet – a négy kandidáns gén jellemzői

Kutatásaink során azonosítottunk négy olyan gént (*EFCAB10*, *GLOD5*, *NONO* és *TMEM70*), amelyeket korábban nem vizsgáltak szarvasmarhában, illetve nem hoztak összefüggésbe tejtermelési tulajdonságokkal.

Az egyes gének jellemzőinek rövid bemutatása és a tejtermelésre gyakorolt hatásuk, hatásmechanizmusuk.

EFCAB10

Szerepe és jellemzői:

- Kalciumion-kötés: Az *EFCAB10* fehérje rendelkezik EF-hand motívumokkal, amelyek képesek a kalciumionok reverzibilis megkötésére.
- Sejtjelátvitel: Bár az *EFCAB10* pontos működése még nem teljesen tisztázott, feltételezik, hogy szabályozó szerepet tölt be bizonyos kalciumfüggő jelátviteli útvonalakban, amelyek befolyásolják a sejtek anyagcseréjét, aktivitását vagy differenciálódását.

Potenciális hatásai szarvasmarhában:

- Mivel a kalciumionok fontosak a tejtermelés, különösen a tejösszetevők szintézise (kazein-fehérjék), valamint az izomműködés (pl. a tőgyürítő izmok működése) szempontjából, az *EFCAB10* gén szerepe is jelentős lehet.
- Metabolikus stabilitás: A kalciumháztartás zavara például szubklinikai tejláz formájában is jelentkezhet, így az *EFCAB10* gén szabályozó hatása ebben is érdekes célpont lehet.
- Tejtermelési tulajdonságok: egyes GWAS-vizsgálatok alapján az *EFCAB10* gén közelében elhelyezkedő SNP-k kapcsolatba hozhatók a tejhozam genetikai hátterével.

Összegzés:

Az *EFCAB10* gén egy kevésbé vizsgált, ám ígéretes célpont, amelynek kalciumkötő tulajdonságai révén valószínűleg szerepe van a tejtermelés szabályozásában, a metabolikus stabilitás fenntartásában és esetleg más, kalciumion-függő sejtes folyamatokban.

GLOD5

Szerepe és jellemzői:

- Fehérjeterméke egy glioxaláz domén jelenlétét mutatja, ami azt sugallja, hogy a gén által kódolt fehérje szerepet játszhat a sejten belüli méregtelenítő folyamatokban, különösen az alfa-oxoaldehydeik, például a metilglioxál semlegesítésében.
- Feltételezhető, hogy a *GLOD5* részt vesz az oxidatív stressz elleni védekezésben, hasonlóan a klasszikus glioxaláz enzimekhez, bár aktivitását vagy pontos enzimikus funkcióját még nem validálták teljeskörűen.

- Az emberi és más emlősgenomokban is megtalálható, de kifejeződésének mintázata, funkcionális aktivitása és szerepe konkrét sejttípusokban vagy szövetekben még további vizsgálatokat igényel.

Összegzés:

A disszertációban azonosított SNP alapján a *GLOD5* lókusznak közelében található variáns összefüggést mutatott tejtermelési tenyésztéssel, ami azt feltételezi, hogy a *GLOD5* gén vagy annak szabályozó régiói hatással lehetnek a metabolikus hatékonyságra, például az energiaforgalomra, aminek közvetett szerepe lehet a tejtermelés genetikai hátterében. Elképzelhető, hogy a *GLOD5* bizonyos metabolikus vagy detoxifikációs útvonalak finomhangolásában vesz részt, befolyásolva ezáltal a termelési teljesítményt, különösen stresszes vagy intenzív termelési körülmények között.

NONO

Szerepe és jellemzői:

- A *NONO* gén egy multifunkcionális nukleáris fehérjét kódol, amely fontos szerepet játszik a transzkripció szabályozásban, az RNS érésében, valamint a DNS-károsodási válaszreakciókban.
- Bár a tejtermelési tulajdonságokkal való közvetlen kapcsolata még kevésbé ismert, a *NONO* gén biológiai szerepe alapján érdekes és potenciálisan releváns célpont lehet a szarvasmarhagenom vizsgálatában is.
- A *NONO* gén által kódolt fehérje RNS- és DNS-kötő képességgel is rendelkezik, és részt vesz:
 - a génexpresszió szabályozásában (transzkripció aktiváció és represszió),
 - az RNS-splicing-, RNS-transzport- és -stabilitási folyamatokban,
 - a sejtmag szerveződésében, különösen a paraspeckle nevű struktúrák kialakításában,
 - valamint a DNS-károsodási válaszreakciókban (például a kettős szálú DNS-törések javításában).
- A disszertációban az SNP-analízis során a *NONO* génhez közeli variáns asszociációt mutatott tejtermelési (különösen a TH- és FH-) tenyésztéssel, ami arra utalhat, hogy a gén környezete hatással van a tejtermelést befolyásoló gének expressziójára.
- A *NONO* gén transzkripció koaktivátorként befolyásolhat olyan géneket, amelyek az anyagcsere-folyamatokhoz, sejtciklushoz vagy tejmirigyfunkcióhoz kapcsolódnak.

- Egyes humán és egérmodellek alapján a *NONO* gén kifejeződése kapcsolatban állhat a hormonális szabályozással is, például az ösztrogénválasz-útvonalakkal, amelyek közvetve a tejtermelő mirigyek fejlődését és működését is befolyásolják.

Összegzés:

A *NONO* gén egy nukleáris multifunkcionális szabályozó fehérjét kódol, amely szerepet játszik a génextpresszió finomhangolásában, az RNS-folyamatokban és a sejtválaszokban. Bár direkt tejtermelési szerepe nem bizonyított, a gén környező régiójában elhelyezkedő SNP-k asszociációja a tenyésztékekkel azt sugallja, hogy a *NONO* gén szabályozási hatásai közvetetten befolyásolhatják a tejtermelés genetikai hátterét.

TMEM70

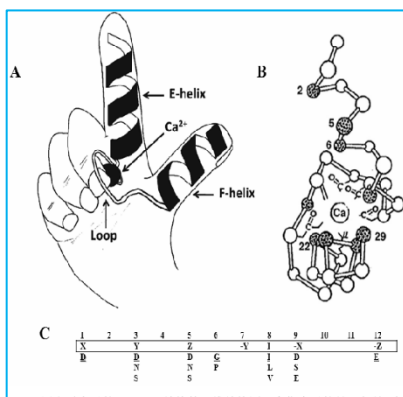
Szerepe és jellemzői:

- A *TMEM70* gén (*Transmembrane Protein 70*) egy mitokondriális transzmembránfehérjét kódoló gén, amely kulcsszerepet játszik a sejt energiatermelésében.
- A gén a mitokondrium belső membránjában működik, és elsősorban az ATP-szintáz (komplex V) megfelelő összeszerelésében és működésében vesz részt.
- Az energiatermelést (ATP-előállítást) támogatja oxidatív foszforiláció révén.

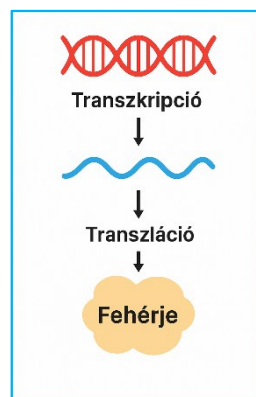
Összegzés:

Bár a szarvasmarhában a *TMEM70* szerepét eddig nem részletezték széles körben, az általunk azonosított SNP-eredmények alapján a gén a tejtermelésben potenciálisan szerepet játszó új célpontként merülhet fel. A mitokondriális működés és az energiatermelés hatékonysága közvetlen összefüggésben lehet a tej szintéziséhez szükséges sejt szintű metabolikus aktivitással, így a termelési hatékonyságot és perzisztenciát is befolyásolhatja. A *TMEM70* jelenléte a 28. kromoszóma nagy hatású SNP-markerei közelében további kutatásokat indokolhat a tejhozam genetikai hátterének pontosítása céljából.

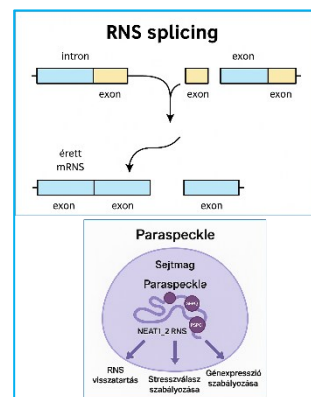
Képi illusztrációk a *EFCAB10*; *NONO* és a *TMEM70* gének funkcióihoz



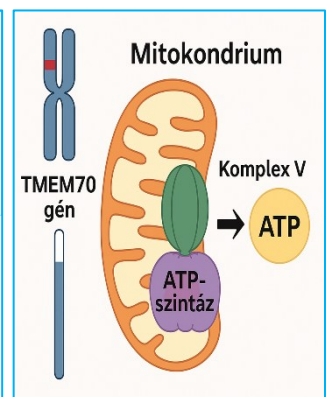
EFCAB10
Forrás: Springer, Dordrecht



NONO gén funkciói
Saját ábra



NONO gén funkciói
Saját ábra



TMEM70 gén funkciói
Saját ábra

TELJES TERJEDELMŰ, EREDETI KÖZLEMÉNYEK

FULL-LENGTH ORIGINAL RESEARCH PAPERS



Article

Different Breeding Values Under Uniform Environmental Condition for Milk Production Yield Traits in Holstein-Friesian Cows

László Bognár, Zsolt Jenő Kőrösi, István Anton, Szabolcs Bene and Ferenc Szabó



Article

Different Breeding Values Under Uniform Environmental Condition for Milk Production Yield Traits in Holstein-Friesian Cows

László Bognár ¹, Zsolt Jenő Kőrösi ¹, István Anton ², Szabolcs Bene ^{2,*} and Ferenc Szabó ³¹ National Association of Hungarian Holstein Friesian Breeders, Lőportár u. 16, H-1134 Budapest, Hungary² Institute of Animal Sciences, Georgikon Campus, Hungarian University of Agriculture and Life Sciences, Deák Ferenc u. 16, H-8360 Keszthely, Hungary³ Department of Animal Sciences, Albert Kázmér Faculty of Agriculture and Food Sciences, Széchenyi István University, Vár t. 2, H-9200 Mosonmagyaróvár, Hungary

* Correspondence: bene.szabolcs.albin@uni-mate.hu

Simple Summary: This study compared the phenotypic performance and breeding values (BV) of Holstein-Friesian cows using three BLUP (best linear unbiased prediction) estimation methods: pedigree (BV_{Pedigree}), traditional (BV_{BLUP}) and genomically enhanced (BV_{Genomic}) to validate the different estimation models. The novelty of this study lies in the validation of BV estimation methods under uniform environmental conditions. It also provides valuable insights into the implementation of genomic selection in a real production environment, providing practical guidance for optimizing genetic improvement strategies.

Abstract: In this study, 1,616,549 Holstein-Friesian females were genotyped for genomic evaluation of genetic merit (BV_{Genomic}). Genotyping was performed using the EuroGenomics MD v3.0 chipset on the Illumina microarray scanner platform operated by an accredited Illumina laboratory. In addition, international and national reference populations were used for traditional BLUP breeding value (BV) estimation for both individuals (BV_{BLUP}) and parents (BV_{Pedigree}). A single-step BLUP animal model was used for this estimation. A sample of 190 first lactation progeny cows from a single herd, reared and kept under consistent environmental conditions, was used to validate the three types of BV estimation methods. Correlation and regression analysis were used to study the association between the phenotypic performance and the results of three different estimation models. The average production of the 305-day standard lactation was 10,910.5 kg milk, 397.86 kg butterfat and 365.33 kg protein. Comparative analyses showed that BV_{BLUP} had the highest accuracy, followed by BV_{Genomic}, while BV_{Pedigree} was the least reliable, R² = 0.37 to 0.48; 0.09 to 0.23; 0.02 to 0.06, respectively.

Keywords: Holstein-Friesian cows; milk; fat and protein production; pedigree BLUP; traditional BLUP; genomically enhanced BLUP



Received: 12 November 2024

Revised: 24 December 2024

Accepted: 26 December 2024

Published: 28 December 2024

Citation: Bognár, L.; Kőrösi, Z.J.; Anton, I.; Bene, S.; Szabó, F. Different Breeding Values Under Uniform Environmental Condition for Milk Production Yield Traits in Holstein-Friesian Cows. *Animals* **2025**, *15*, 51. <https://doi.org/10.3390/ani15010051>

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breeding value (BV) is defined as the genetic potential of a breeding animal as a parent [1]. BV can be estimated from various sources of data on the animal itself, its genetically similar collateral relatives and/or the performance of its offspring. The Best Linear Unbiased Prediction (BLUP) method has been developed to consider performance data from genetically diverse contemporary groups, such as those from different farms. This method is considered unbiased because it incorporates more extensive data in subsequent

predictions for the same animal. Various versions of BLUP are now widely used as an advanced statistical model that evaluates all animals in a population.

Subsequently, a genomic BV estimation (BV_{Genomic}) method was developed that incorporates genomic data and SNP (single nucleotide polymorphism) information from DNA. This method is based on both the SNP information of an individual and the relationship between SNP information and performance, deregressed from the BLUP data of the reference population [2]. Wiggans et al. [3] discussed the impact of genomic selection on dairy breeding, noting that SNP genotyping has allowed faster genetic progress by reducing the generation interval since young animals or embryos can be genotyped. Two main methods (Bayesian and BLUP) have been extensively studied and applied [4]. Echeverri et al. [5] found that predicting BVs using BLUP, MBLUP and Bayes C in the Holstein-Friesian breed produced different results in terms of the magnitude of the estimated values, although BVs based on animal rankings showed no significant differences. Abaci et al. [6] found that among the different methods, the correlation was highest between BLUP and Bayes C_{pi}, while the correlation was lowest between BLUP and Bayes A. Despite their complexity, Bayesian methods are less widely used in practical breeding than BLUP 79 methods, as highlighted by Wang et al. [4].

Koivula et al. [7] have compared different estimation methods and found that SNP-BLUP and G-BLUP yielded the same validation reliability, while H-BLUP (H-matrix best linear unbiased prediction) provided slightly higher reliability, indicating a marginal advantage of using H-BLUP for genomic evaluation. Lee et al. [8] observed that the reliability of genomic estimated BVs (GEBVs) increased by an average of 9% over traditional estimated BVs (EBVs), with a 7% increase in cows with test records and about a 4% increase in bulls with progeny records. Herrera et al. [9] have evaluated the accuracy of BV estimations using genomic best linear unbiased prediction (GBLUP) and single-step GBLUP (ssGBLUP) compared to Pedigree BLUP (pBLUP) in Philippine dairy buffaloes, finding that genomic methods provided more accurate predictions than pBLUP. Massender et al. [10] studied two goat breeds and observed a small gain in validation accuracy for GEBV relative to pedigree-based EBV in the Alpine breed but not in the Saanen breed, potentially due to limitations in the validation design. Abdel-Shafy et al. [11] reported that using a genotyping array with 90K markers, the prediction accuracy of 0.61 appears suitable under the current challenges. Zhang et al. [12] compared pedigree BLUP, GBLUP, and ssGBLUP, finding that the reliability of EBV can be enhanced by the use of GEBV.

Hayes et al. [13] have applied GBLUP to a Holstein-Friesian reference population and found that while realized and expected accuracies were reasonably aligned, the expected accuracies did tend to over-predict the realized accuracies by an average of 8% across various traits. Echeverri et al. [5] observed that Spearman correlation coefficients between BVs obtained by different methods exceeded 0.5. The results indicated that while the magnitudes of the BVs using BLUP, Modified BLUP (MBLUP), and Bayes C varied across these methods, the rankings of animals based on their BVs did not differ significantly, and despite methodological differences, the relative genetic merit assigned to each animal remained consistent. Cesarani et al. [14] have reported that the inclusion of genomic information could enhance the accuracy of BV estimations and accelerate genetic progress for milk ability in Italian Simmental cattle. Aguilar et al. [15] evaluated four genetic analysis methods for U.S. Holstein cattle, comparing pedigree-only, combined pedigree-genomic (single-step), full data with pedigree, and multi-step approaches. The single-step method, which integrates both pedigree and genomic information, produced genomic predictions with accuracy and bias similar to multi-step methods and could adapt to various population structures. This approach is expected to become increasingly advantageous as more animals are pre-selected based on genotypes.

Fessenden et al. [16] emphasized the importance of selection indices in breeding programs, particularly for predicting an animal's genetic potential in terms of economic merit. Their study retrospectively examined the effectiveness of a specific selection index, which included genomically-enhanced predicted transmitting abilities, to determine its capacity to forecast observed lifetime profit in U.S. Holstein cattle. The observation is that selection indices are fundamental in breeding programs, introduced in the 1940s to promote balanced genetic gains across traits influencing productivity and economic outcomes, according to Hazel [17]. These indices combine multiple trait data into a single value, helping to rank animals and guide breeding decisions [18]. Typically, selection indices estimate the genetic potential for overall economic merit [19]. Commercial genomic tests commonly incorporate selection indices, enabling dairy producers to evaluate heifers for strategic culling and breeding [20]. Lourenco et al. [21] found that incorporating genomic data in livestock breeding, particularly in dairy cattle, enhances the accuracy of estimated BVs (EBVs) due to improved insights into relationships and linkage disequilibrium (LD) with quantitative trait loci (QTLs). They noted that while adding female genotypes to a primarily male training set offers minimal accuracy gains for young bull evaluations, genotyping females remains valuable for intra-herd selection and identifying elite dams. Genomic selection thus enables early selection and intensifies genetic progress, particularly in high-selection pathways.

However, genomic BV estimation and selection have been used in cattle breeding for several years, and the permanent control of the applied methods may be beneficial, not only from a scientific but also from a practical point of view.

Although numerous publications exist, studies systematically comparing and evaluating various methods of BV estimation under uniform conditions remain limited, and accessible results derived from a substantial number of cows maintained within the same herd are conspicuously lacking.

The objective of this study was to estimate BVs using three different methods, such as the $BV_{Pedigree}$, BV_{BLUP} and $BV_{Genomic}$, to compare the most significant productive traits of Holstein-Friesian cows in Hungary. The accuracy, reliability, and usability of the three BV estimation methods were evaluated for first-lactation cows of similar age born and managed within the same herd under uniform conditions.

The novelty of this study lies in the validation of BVs by directly comparing different estimation methods with actual realized performance in uniform environmental conditions, providing new insights into the practical application of genomic selection in Holstein cattle breeding.

2. Materials and Methods

This study is based on three types of BV data of 1,616,549 Holstein-Friesian females validated by the phenotypic performance of 190 first lactation progeny cows.

2.1. The Sample Population Database and the Estimated Traits

The 1,616,549 females used for BV estimation were kept in different herds on different large-scale dairy farms in Hungary, with an average of 453 milk-recorded, herd-book-registered cows per herd [22]. The housing system employed a loose-housing free-stall barn design, featuring either a common lying area or an open lounging configuration. Milking is usually performed in a milking parlor or on some farms by robotic milking.

The cows were fed a TMR (Total Mixed Ration) based system throughout the year. The ration consisted mainly of maize silage or silage of other cereals and concentrates of cereals and protein sources, supplemented with minerals and vitamins. The proportion of the cows' daily ration was based on their milk production, lactation or dry period stage.

The 190 first lactation progeny cows used to validate the different BVs were born in the same year between March and September 2018. They were reared in the same herd and kept on the same farm, which is one of the largest commercial Holstein dairies in the country with 1000 milking, milk-recorded cows. This controlled housing and feeding regime was essential for minimizing environmental variation, ensuring that the production traits and genetic evaluations reflected inherent genetic differences.

The 190 cows calved and started their first lactation. After 305 days, complete production and type classification performance data were obtained. The production traits measured were 305-day milk production (MLK, kg), fat production (FAT, kg) and protein production (PRO, kg).

Three types of BV, pedigree (BV_{Pedigree}), traditional (BV_{BLUP}) and genomically enhanced (BV_{Genomic}), were available for all females.

The processing of the performance and BV data was performed according to the method described by Stoop et al. [23].

2.2. Breeding Value Estimation Methods

The details of the aforementioned three kinds of BV estimation methods are as below.

The BV_{Pedigree} was calculated as a simple mean value of the BV_{BLUP} of the dam and that of the sire as follows:

$$BV_{\text{Pedigree}} = (BV_{\text{BLUPd}} + BV_{\text{BLUPs}})/2 \quad (1)$$

(where: BV_{Pedigree} = pedigree BV of the cow; BV_{BLUPd} = BV_{BLUP} of the dam; BV_{BLUPs} = BV_{BLUP} of the sire.)

The BV_{BLUP} : Using the BLUP model, two matrices were created. One of these was the database matrix, and the other was the pedigree matrix. The pedigree matrix of relatives included pedigree data for full sibs, half-sibs, sires, dams, and grandparents. BLUP models contained information for maternal genetic effects and maternal permanent environmental effects as random effects. The models were constructed as follows:

$$y = X_b + Z_a + W_{pe} + e \quad (2)$$

(where: “y” is the vector of observations; “b” is the vector of fixed effects; “a” is the vector of random animal effects; “pe” is the random vector of permanent environmental effects; “e” is the vector of random residual effects; X, Z and W are the incidence matrices relating records to fixed, animal and random permanent environmental effects, respectively)

The fixed effects were a herd, year, season, parity and age. Random effects include genetic and environmental influences specific to individual animals but not systematically attributable to the fixed effects.

The BV_{Genomic} was estimated using the following: The SNP effect for BV_{Genomic} was estimated from BV_{BLUP} and its reliability, as well as from the genotype. The derived de-regressed BV (de-regressed proof, DRP) from the BV_{BLUP} was used for the calculation of the direct genomic BV (DGV), according to Van Raden et al. [24].

$$DRP = PA + (EBV - PA) \times (EDC_{\text{parents} + \text{progeny}} / EDC_{\text{progeny}}) \quad (3)$$

The DGV is based on the Bayesian multi-QTL model of Meuwissen and Goddard [25], where the effects of dense SNPs across the whole genome are fitted directly without the use of haplotypes or identical-by-descent probabilities [26]. Although the method can be

applied for multiple traits simultaneously, the routine genomic evaluations are single trait analyses, i.e., $m = 1$. For m traits, the model is [23]:

$$y_i = \mu + u_i + \sum_{j=1}^{40,947} z_{ij}q_jv_j + e_i \quad (4)$$

(where: y_i = vector of phenotypes (deregressed proofs) of bull i ; μ = vector of fixed trait means; u_i = vector of random polygenic effects of bull i ; q_j = vector of random non-scaled SNP effects for SNP j with alleles 0, 1, and 2, where SNP allele 0 corresponds to a missing genotype; v_j = random scaling vector for SNP j ; z_{ij} = design vector for bull i and SNP j — $z_{ij} = [0 \ 2 \ 0]$, $[0 \ 1 \ 1]$, $[0 \ 0 \ 2]$ or $[2 \ 0 \ 0]$ for homozygous (AA), heterozygous (AB), homozygous (BB), or non-genotyped (00) bulls at SNP j , respectively; e_i = vector of residuals of bull i .)

2.3. Correlation and Regression Analysis

To evaluate the normality of the production trait data, the Kolmogorov–Smirnov test was employed, while Levene’s test was used to assess the homogeneity of variances. A multifactor analysis of variance (ANOVA) for the mentioned traits was conducted. Pearson correlation coefficients were calculated between different BVs (r_g), moreover between BVs and phenotype (r_{gp}), as well as Spearman rank correlation between different BVs (r_s). Linear regression was employed to evaluate the relationships among phenotypic traits and BVs, as outlined by Turney [27]. Throughout the statistical analysis, the methodologies proposed by Seo et al. [28] were followed. Phenotypic records and EBVs were standardized to z-scores before calculating correlations.

2.4. Used Softwares

The data were prepared using Microsoft Excel 2019 and Microsoft Word 2019. The evaluation of the database was carried out using the statistical software package SPSS version 27.0 [29].

3. Results

The MLK, FAT and PRO of the studied population were quite favorable, as shown in Table 1. The group of 190 cows was relatively homogeneous in terms of production traits, indicated by the coefficient of variation (CV%) being less than 15%.

Table 1. The basic statistics of the sampled cows.

Trait	MLK (kg)	FAT (kg)	PRO (kg)
Number of cows (n)	190	190	190
Mean value	10,910.50	397.86	365.33
SD	1453.70	45.24	39.80
CV%	13.32	11.37	10.90
SE	105.46	3.28	2.89
Min	6505	260	196
Max	13781	511	451

MLK = 305-day milk production; FAT = 305-day fat production; PRO = 305-day protein production.

Data in Table 2 reveal a significant range between the minimum and maximum values of the different BVs; however, the standard errors (SE) relative to the mean values are relatively low (below 5%). In all instances, the data for $BV_{Pedigree}$ are higher than those for the other two types of BVs. This discrepancy is attributed to the stepwise base change in BV

estimation carried out every five years between the parental generation and the offspring cow generation.

Table 2. Basic statistics of the analyzed breeding values.

Trait/Breeding Value	N	Mean	SE	Min	Max
MLK (kg)					
BV _{BLUP}	190	739.16	31.55	−397	1779
BV _{Genomic}	190	718.11	32.31	−357	2052
BV _{Pedigree}	170	1379.57	29.98	−186	2354
FAT (kg)					
BV _{BLUP}	190	34.47	1.00	0	65
BV _{Genomic}	190	33.44	1.24	−24	75
BV _{Pedigree}	170	55.58	0.97	22	87
PRO (kg)					
BV _{BLUP}	190	27.01	0.78	−2	56
BV _{Genomic}	190	26.05	0.99	−18	66
BV _{Pedigree}	170	49.33	0.91	6	78

MLK = 305-day milk production; FAT = 305-day fat production; PRO = 305-day protein production; BV_{BLUP} = traditional BLUP animal model; BV_{Genomic} = genomically enhanced BLUP animal model; BV_{Pedigree} = pedigree BLUP model.

Table 3 summarizes the correlation coefficients between BVs and phenotypic performances and that of BVs estimated in three different ways.

Table 3. Phenotypic, genetic and rank correlations between breeding values and phenotypic traits.

Correlations (r)	BV _{BLUP}	BV _{Genomic}	BV _{Pedigree}
MILK	$r_{gp} = 0.70; p < 0.01$	$r_{gp} = 0.48; p < 0.01$	$r_{gp} = 0.24; p < 0.01$
BV _{BLUP}		$r_g = 0.67; p < 0.01$	$r_g = 0.66; p < 0.01$
BV _{Genomic}		$r_s = 0.65; p < 0.01$	$r_s = 0.50; p < 0.01$
			$r_g = 0.53; p < 0.01$
			$r_s = 0.40; p < 0.01$
FAT	$r_{gp} = 0.69; p < 0.01$	$r_{gp} = 0.32; p < 0.01$	$r_{gp} = 0.15; NS$
BV _{BLUP}		$r_g = 0.67; p < 0.01$	$r_g = 0.56; p < 0.01$
BV _{Genomic}		$r_s = 0.65; p < 0.01$	$r_s = 0.57; p < 0.01$
			$r_g = 0.43; p < 0.01$
			$r_s = 0.41; p < 0.01$
PRO	$r_{gp} = 0.61; p < 0.01$	$r_{gp} = 0.31; p < 0.01$	$r_{gp} = 0.15; p < 0.05$
BV _{BLUP}		$r_g = 0.66; p < 0.01$	$r_g = 0.60; p < 0.01$
BV _{Genomic}		$r_s = 0.66; p < 0.01$	$r_s = 0.12; NS$
			$r_g = 0.56; p < 0.01$
			$r_s = 0.56; p < 0.01$

BV_{BLUP} = traditional BLUP animal model; BV_{Genomic} = genomically enhanced BLUP animal model; BV_{Pedigree} = pedigree BLUP model; r_{gp} = correlation between the phenotype and genotype; r_g = genetic correlation; r_s = rank correlation.

The values of correlation coefficients are positive in each case, and most of them are significant. As is seen in the table, the BV_{BLUP} shows a stronger relationship with the phenotypic performances ($r_{gp} = 0.61$ – 0.70) than BV_{Genomic} ($r_{gp} = 0.31$ – 0.48). The weakest association ($r_{gp} = 0.15$ – 0.24) between genotype and phenotype was found in the case of BV_{Pedigree}. The BV_{Genomic} showed moderate and strong association ($r_g = 0.66$ – 0.67) with BV_{BLUP}. There were also a moderate and strong association ($r_g = 0.56$ – 0.66) between BV_{BLUP} and BV_{Pedigree} and a little bit lower, moderate or strong relationship ($r_g = 0.43$ – 0.56) between BV_{Genomic} and BV_{Pedigree}.

The rank correlation values ($r_g = 0.66$ – 0.67) between BV_{BLUP} and $BV_{Genomic}$ indicate a strong ($r_s = 0.65$ – 0.66) association. The correlation coefficients ($r_s = 0.12$ – 0.57) between the BV_{BLUP} and the $BV_{Pedigree}$ and between $BV_{Genomic}$ and $BV_{Pedigree}$ also show a moderate association ($r_s = 0.40$ – 0.56).

The results of regression analyses for BVs against phenotypic traits are presented in Table 4. Significant associations were found at a high level ($p < 0.01$), except for the $BV_{Pedigree}$ for FAT. All regression coefficient (b) values were positive, with the highest observed for BV_{BLUP} ($b = 0.17$ – 0.21), followed by $BV_{Genomic}$ ($b = 0.10$ – 0.15) and the lowest for $BV_{Pedigree}$ ($b = 0.04$ – 0.07). The data clearly demonstrate that the highest determination ($R^2 = 0.37$ – 0.48) for each phenotypic trait was achieved by BV_{BLUP} , with $BV_{Genomic}$ showing a lower determination ($R^2 = 0.09$ – 0.23) and $BV_{Pedigree}$ the least ($R^2 = 0.02$ – 0.06). Among the production traits, MLK and FAT were more strongly determined by BVs than PRO. Slope values ($b = 0.17$ – 0.21) were positive in each case, being highest for BV_{BLUP} , lower for $BV_{Genomic}$ (0.10 – 0.15), and lowest for $BV_{Pedigree}$ (0.04 – 0.07).

Table 4. Regression analysis results.

Breeding value (Y)	Trait (X)	Slope			Intercept			Fitting	
		b	SE	p	a	SE	p	R ²	p
BV_{BLUP}	MLK	0.21	0.02	<0.01	−1529.56	172.66	<0.01	0.48	<0.01
$BV_{Genomic}$	MLK	0.15	0.02	<0.01	−879.32	216.02	<0.01	0.23	<0.01
$BV_{Pedigree}$	MLK	0.07	0.02	<0.01	665.15	229.32	<0.01	0.06	<0.01
BV_{BLUP}	FAT	0.21	0.02	<0.01	−48.80	6.51	<0.01	0.47	<0.01
$BV_{Genomic}$	FAT	0.12	0.03	<0.01	−14.20	10.47	<0.01	0.10	<0.01
$BV_{Pedigree}$	FAT	0.04	0.02	NS	38.83	8.66	<0.01	0.02	NS
BV_{BLUP}	PRO	0.17	0.02	<0.01	−33.12	5.75	<0.01	0.37	<0.01
$BV_{Genomic}$	PRO	0.10	0.02	<0.01	−12.12	8.73	<0.01	0.09	<0.01
$BV_{Pedigree}$	PRO	0.05	0.02	<0.05	32.06	8.77	<0.01	0.02	<0.05

MLK = 305-day milk production; FAT = 305-day fat production; PRO = 305-day protein production; BV_{BLUP} = traditional BLUP animal model; $BV_{Genomic}$ = genomically enhanced BLUP animal model; $BV_{Pedigree}$ = pedigree BLUP model.

According to both the correlation and regression results, the BV_{BLUP} method was found to be the most reliable, followed by the $BV_{Genomic}$, and the $BV_{Pedigree}$ method showed the least reliability in the same environmental context.

4. Discussion

BV estimation presents significant challenges, such as accurately representing genetic effects, minimizing environmental influences, and enhancing the heritability of specific traits to improve estimation accuracy. Two primary strategies can enhance the precision of predictions: increasing the number of relatives contributing data for evaluation and minimizing environmental effects as much as possible [1].

Large-scale genetic evaluations using BLUP techniques can manage data from multiple herds across a country or even broader geographical areas. Furthermore, genomically enhanced estimation methods facilitate more precise determination of genotypes and allow for the estimation of BVs at earlier ages of the animals. Schefers et al. [30] highlighted that genomic selection enhances genetic progress in dairy cattle by improving the accuracy of genetic predictions for young animals, shortening the generation interval, and increasing selection intensity. These advancements nearly double genetic gains in economically valuable traits. Breeders now routinely use genomically enhanced BVs (GEBVs) to select animals for reproduction, while AI companies employ genomic testing to identify elite bulls and females. This approach has become integral to modern dairy breeding and is

expected to remain essential in the industry for future progress. Štrbac et al. [31] evaluated six univariate and two multivariate BLUP models for BV estimation in Holstein cattle, comparing different combinations of the numerator relationship matrix (NRM) and genomic relationship matrix (GRM). Results showed that a combination of NRM and GRM in univariate analysis achieved higher accuracy than NRM alone, while multivariate models with repeated measures gave the highest accuracy for all animals. For genotyped animals, ssGBLUPp provided the most accurate BV estimates. The study suggests transitioning to multivariate, repeated measures until a robust reference population is established.

In our model study, which compared different BV estimation methods, we made efforts to minimize environmental impacts on dairy cow performance.

The results in the same cases are similar; some other cases are slightly different from the data published in the literature.

In this study, positive correlations were consistently observed both between different BVs and between phenotypic performances and BVs. This finding partially aligns with and partially diverges from the results of Echeverri et al. [5], who reported that Spearman correlation coefficients between BVs obtained by different methods exceeded 0.5, whereas linear regression coefficients varied between -2.10 and 1.58 .

The correlation coefficients between genotype and phenotype, indicated by BV_{BLUP} , suggest a high degree of genetic determination ($r_{gp} = 0.61-0.71$), as further supported by large values ($R^2 = 0.37-0.48$) of the coefficient of determination. $BV_{Genomic}$ demonstrated medium to high genetic determination ($r_{gp} = 0.31-0.62$, $R^2 = 0.09-0.38$), albeit less than that observed with BV_{BLUP} . $BV_{Pedigree}$ showed the lowest correlation ($r_{gp} = 0.15-0.29$) and determination coefficients ($R^2 = 0.02-0.08$), indicating a weaker predictive ability for the outcome phenotype compared to the other models.

Considering these correlation and determination coefficients as indicators of accuracy and reliability, the results of our study suggest that BV_{BLUP} is the most accurate and reliable method, followed by $BV_{Genomic}$ and finally by $BV_{Pedigree}$.

Our findings align with those of Echeverri et al. [5] and Abaci et al. [6], who concluded that the prediction of BVs using BLUP, MBLUP and Bayes C displayed differences in terms of magnitude from the estimated values. Similarly, our results are somewhat in line with those of Herrera et al. [9], who reported that genomic methods genomic BLUP and single-step genomic BLUP (GBLUP and ssGBLUP) provide more accurate predictions than $BV_{Pedigree}$, with average accuracies for GBLUP and ssGBLUP at 0.24 and 0.29, respectively, compared to 0.21 and 0.22 for $BV_{Pedigree}$. However, our findings diverge from those of Koivula et al. [7], Lee et al. [8] and Cesarani et al. [14], who observed an increase in the reliability of $BV_{Genomic}$ over BV_{BLUP} .

Further, our results slightly differ from those of Zhang et al. [12], who compared $BV_{Pedigree}$ with GBLUP and ssGBLUP, noting that the reliability of estimated BVs could be improved from 0.9% to 3.6%, and the reliability of $BV_{Genomic}$ for the genotyped population could reach 83%.

The variance between our results and those documented in the literature likely stems from the specific conditions of our study's herd, which was exposed to minimal environmental impacts. This reduced environmental variability might have diminished the observable advantages of $BV_{Genomic}$ estimation methods compared to the BV_{BLUP} approach.

The observation that the $BV_{Pedigree}$ values are consistently higher than the other two types of BVs (BV_{BLUP} and $BV_{Genomic}$) while BV_{BLUP} shows the strongest correlation can be explained by considering the following factors:

Simplified calculation of $BV_{Pedigree}$: $BV_{Pedigree}$ is based solely on the average of parental BVs without accounting for additional genetic or environmental factors. This

method does not incorporate phenotypic or genotypic data specific to the individual, leading to an upward bias, especially if the parents' evaluations are themselves overestimated.

Inclusion of phenotypic data in BV_{BLUP} : BV_{BLUP} incorporates phenotypic performance data, adjusting for environmental effects and providing a more individualized estimate of genetic merit. This integration enhances the precision and reliability of the predictions, which likely explains its stronger correlation with actual performance.

Impact of genomic information in $BV_{Genomic}$: $BV_{Genomic}$ uses SNP data to enhance accuracy but depends on the size and structure of the reference population. In cases where the reference population is small or not representative, the genomic prediction may have lower accuracy, potentially leading to weaker correlations compared to BV_{BLUP} .

Greater variability in $BV_{Pedigree}$ estimates: Since $BV_{Pedigree}$ lacks adjustments for individual-specific data, it may exhibit greater variability compared to BV_{BLUP} or $BV_{Genomic}$. This could explain why it is consistently higher but less strongly correlated with actual performance.

In uniform environmental conditions, as indicated in the study, BV_{BLUP} may better capture the genetic contribution to traits due to its reliance on phenotypic data, resulting in stronger correlations compared to $BV_{Pedigree}$ or $BV_{Genomic}$.

We hope that by integrating BV results into herd management strategies, producers enhance the overall efficiency and profitability of their operations, demonstrating the practical utility and impact of the genetic evaluation procedures validated in this study.

5. Conclusions

This study demonstrated significant differences in the accuracy and reliability of BVs estimated using $BV_{Pedigree}$, BV_{BLUP} and $BV_{Genomic}$ under identical environmental conditions in Hungary. BV_{BLUP} exhibited the highest reliability, followed by $BV_{Genomic}$ and then $BV_{Pedigree}$.

Our finding does not diminish the relevance and applicability of the $BV_{Genomic}$ since it provides critical time-saving advantages, offering earlier predictions for traits such as management, longevity and lifetime performance. It also allows farmers to optimize herd size and reduce maintenance costs. However, BV_{BLUP} may serve as a complementary control, ensuring robustness in genetic evaluations where feasible.

The results obtained in this study highlight the importance of tailoring breeding strategies to specific environmental and population contexts and may improve the accuracy and application of BV estimation methods in dairy production for female selection.

In addition to all of this, in the future, it would be worthwhile to perform the calculations repeatedly on a larger database using mixed mathematical methods (e.g., transformed data). The results obtained on a larger database could provide an opportunity to draw further conclusions.

Author Contributions: Conceptualization, L.B., F.S. and Z.J.K.; methodology, L.B., F.S. and I.A.; software, S.B.; validation, L.B., F.S. and I.A.; formal analysis, S.B. and I.A.; investigation, L.B. and Z.J.K.; resources, L.B., F.S. and Z.J.K.; data curation, S.B., F.S. and L.B.; writing-original draft preparation, L.B., F.S., Z.J.K. and I.A.; writing-review and editing, L.B., F.S. and Z.J.K.; visualization, Not applicable; supervision, L.B., F.S. and I.A.; project administration, F.S., S.B. and L.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research and publication received no external funding.

Institutional Review Board Statement: This study did not require approval from the Ethical Committee on Animal Experiments, since only data processing was performed.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data for this study were collected and generated by the National Association of Hungarian Holstein Friesian Breeders. These data can be reviewed and checked at their facilities. For more information or to request access to the data, please visit their website or contact them directly: Website: www.holstein.hu; Email: bognar@holstein.hu; Postal Address: Lóportár Street 16, H-1134 Budapest, Hungary.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

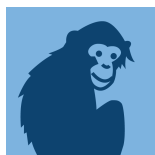
MLK	305-day milk production
FAT	305-day milk fat production
PRO	305-day milk protein production
BLUP	Best Linear Unbiased Prediction
MBLUP	Modified BLUP
BV	Breeding value
BV _{BLUP}	BV estimated with BLUP animal model method
BV _{Genomic}	BLUP animal model enhanced with BV estimated from effects of single nucleotide polymorphisms (SNP)
BV _{Pedigree}	BLUP animal model BV calculated from pedigree (parents) BV
EBV	estimated conventional Hungarian BV
EDC	estimated daughter contributions.
PA	parent average (no female EBV is present, this resembles a sire pedigree index)
DRP	de-regressed proof/de-regressed BV from the BV _{BLUP}
DGV	direct genomic BV

References

1. Bourdon, R.M. *Understanding Animal Breeding*; Prentice-Hall International: Upper Saddle River, NJ, USA, 2000.
2. Calus, M.P.L. Genomic breeding value prediction: Methods and procedures. *Animal* **2010**, *4*, 157–164. [[CrossRef](#)] [[PubMed](#)]
3. Wiggans, G.R.; Cole, J.B.; Hubbard, S.M.; Sonstegard, T.S. Genomic selection in dairy cattle: The USDA Experience. *Annu. Rev. Anim. Biosci.* **2017**, *5*, 309–327. [[CrossRef](#)] [[PubMed](#)]
4. Wang, C.; Ding, X.; Liu, J.; Zhang, Q. Bayesian methods for genomic breeding value estimation. *Yi Chuan* **2014**, *36*, 111–118. [[CrossRef](#)]
5. Echeverri, J.; Zambran, J.C.; Herrera, A.L. Genomic evaluation of Holstein cattle in Antioquia (Colombia). *Rev. Colomb. Cienc. Pecu.* **2014**, *27*, 306–314. [[CrossRef](#)]
6. Abaci, S.H.; Önder, H.; Spurlock, D.M. Determining the relationships between genomic and phenotypic breeding values. *Sci. Pap. Ser. D Anim. Sci.* **2016**, *54*, 11–13.
7. Koivula, M.; Strandén, I.; Su, G.; Mäntysaari, E.A. Different methods to calculate genomic predictions—Comparisons of BLUP at the single nucleotide polymorphism level (SNP-BLUP), BLUP at the individual level (G-BLUP), and the one-step approach (H-BLUP). *J. Dairy Sci.* **2012**, *95*, 4065–4073. [[CrossRef](#)]
8. Lee, Y.M.; Dang, C.G.; Alam, M.Z.; Kim, Y.S.; Cho, K.H.; Park, K.D.; Kim, J.J. The effectiveness of genomic selection for milk production traits of Holstein dairy cattle. *Asian-Austral. J. Anim. Sci.* **2020**, *33*, 382–389. [[CrossRef](#)] [[PubMed](#)]
9. Herrera, J.R.V.; Flores, E.B.; Duijvesteijn, N.; Moghaddar, N.; van der Werf, J.H. Accuracy of genomic prediction for milk production traits in Philippine dairy buffaloes. *Front. Genet.* **2021**, *12*, 682576. [[CrossRef](#)] [[PubMed](#)]
10. Massender, E.; Brito, L.F.; Maignel, L.; Oliveira, H.R.; Jafarikia, M.; Baes, C.F.; Sullivan, B.; Schenkel, F.S. Single-step genomic evaluation of milk production traits in Canadian Alpine and Saanen dairy goats. *J. Dairy Sci.* **2022**, *105*, 2393–2407. [[CrossRef](#)]
11. Abdel-Shafy, H.; Awad, M.A.A.; El-Regalaty, H.; Ismael, A.; El-Assal, S.D.S.M.; Abou-Bakr, S. A single-step genomic evaluation for milk production in Egyptian buffalo. *Liv. Sci.* **2020**, *234*, 103977. [[CrossRef](#)]
12. Zhang, J.; Wang, J.; Li, Q.; Wang, Q.; Wen, J.; Zhao, G. Comparison of the efficiency of BLUP and GBLUP in genomic prediction of immune traits in chickens. *Animals* **2022**, *10*, 419. [[CrossRef](#)]
13. Hayes, B.J.; Bowman, P.J.; Chamberlain, A.C.; Verbyla, K.; Goddard, M.E. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* **2009**, *41*, 51. [[CrossRef](#)] [[PubMed](#)]
14. Cesarani, A.; Garcia, A.; Hidalgo, J.; Degano, L.; Vicario, D.; Macciotta, N.P.P.; Lourenco, D. Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. *J. Dairy Sci.* **2021**, *104*, 5719–5727. [[CrossRef](#)]

15. Aguilar, I.; Misztal, I.; Johnson, D.L.; Legarra, A.; Tsuruta, S.; Lawlor, T.J. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* **2010**, *93*, 743–752. [CrossRef]
16. Fessenden, B.; Weigel, D.J.; Osterstock, J.; Galligan, D.T.; Croce, F.D. Validation of genomic predictions for a lifetime merit selection index for the US dairy industry. *J. Dairy Sci.* **2020**, *103*, 10414–10428. [CrossRef] [PubMed]
17. Hazel, L.N. The genetic basis for constructing selection indexes. *Genetics* **1943**, *28*, 476–490. [CrossRef] [PubMed]
18. Cole, J.B.; VanRaden, P.M. Symposium review: Possibilities in an age of genomics: The future of selection indices. *J. Dairy Sci.* **2018**, *101*, 3686–3701. [CrossRef] [PubMed]
19. Shook, G.E. Major advances in determining appropriate selection goals. *J. Dairy Sci.* **2006**, *89*, 1349–1361. [CrossRef] [PubMed]
20. Weigel, K.A.; Hoffman, P.C.; Herring, W.; Lawlor, T.J., Jr. Potential gains in lifetime net merit from genomic testing of cows, heifers, and calves on commercial dairy farms. *J. Dairy Sci.* **2012**, *95*, 2215–2225. [CrossRef]
21. Lourenco, D.A.L.; Fragomeni, B.O.; Tsuruta, S.; Aguilar, I.; Zumbach, B.; Hawken, R.; Legarra, A.; Misztal, I. Accuracy of estimated breeding values with genomic information on males, females, or both: An example on broiler chicken. *Genet. Sel. Evol.* **2015**, *47*, 47–56. [CrossRef]
22. Livestock Performance Testing Ltd., Newsletter. 2023. Available online: <http://static.atkft.hu/pthl/ujsag2312.pdf> (accessed on 3 August 2024).
23. Stoop, W.M.; Eding, H.; Schrooten, C. *Method of Genomic Breeding Value Estimation, Specified for Hungarian Data*; Proposal for the National Association of Hungarian Holstein Friesian Breeders: Budapest, Hungary, 2017; working document.
24. Van Raden, P.M.; VanTassell, C.P.; Wiggans, G.R.; Sonstegard, T.S.; Schnabel, R.D.; Taylor, J.F.; Schenkel, F.S. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* **2009**, *92*, 16–24. [CrossRef] [PubMed]
25. Meuwissen, T.H.E.; Goddard, M.E. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genet. Sel. Evol.* **2004**, *36*, 261. [CrossRef] [PubMed]
26. Calus, M.P.L.; Meuwissen, T.H.E.; de Roos, A.P.W.; Veerkamp, R.F. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* **2008**, *178*, 553–561. [CrossRef]
27. Turney, S. Coefficient of Determination (R²)—Calculation & Interpretation. *Scribbr*. 2023. Available online: <https://www.scribbr.com/statistics/coefficient-of-determination/> (accessed on 23 October 2024).
28. Seo, S.; Jeon, S.; Ha, J.K. Guidelines for experimental design and statistical analyses in animal studies submitted for publication in the Asian-Australasian Journal of Animal Sciences. *Asian-Austral. J. Anim. Sci.* **2018**, *31*, 1381–1386. [CrossRef]
29. IBM Corporation. *IBM SPSS Statistics for Windows, Version 27.0*; IBM Corporation: Armonk, NY, USA, 2020.
30. Schefers, J.M.; Weigel, K.A. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Anim. Front.* **2012**, *2*, 4–9. [CrossRef]
31. Štrbac, L.; Pracner, D.; Šaran, M.; Janković, D.; Trivunović, S.; Ivković, M.; Tarjan, L.; Dedović, N. Mathematical modeling and software tools for breeding value estimation based on phenotypic, pedigree and genomic information of Holstein Friesian cattle in Serbia. *Animals* **2023**, *13*, 597. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

Simultaneous Effects of Single-Nucleotide Polymorphisms on the Estimated Breeding Value of Milk, Fat, and Protein Yield of Holstein Friesian Cows in Hungary

László Bognár, Zsolt Jenő Kőrösi, Szabolcs Albin Bene, Ferenc Szabó, István Anton and Attila Zsolnai

Topic Collection

Applications of Quantitative Genetics in Livestock Production

Edited by

Prof. Dr. Michael E. Davis



Article

Simultaneous Effects of Single-Nucleotide Polymorphisms on the Estimated Breeding Value of Milk, Fat, and Protein Yield of Holstein Friesian Cows in Hungary

László Bognár ^{1,2} , Zsolt Jenő Kőrösi ^{1,2}, Szabolcs Albin Bene ³ , Ferenc Szabó ², István Anton ³ 
and Attila Zsolnai ^{3,*} 

¹ Association of Hungarian Holstein Breeders, Lőportár utca 16., H-1134 Budapest, Hungary; bognar@holstein.hu (L.B.); korosi@holstein.hu (Z.J.K.)

² Albert Kázmér Faculty of Mosonmagyaróvár, Széchenyi István University, Vár tér 2., H-9200 Mosonmagyaróvár, Hungary; szabo.ferenc@sze.hu

³ Institute of Animal Husbandry Sciences, Hungarian University of Agriculture and Life Sciences, Guba Sándor utca 40., H-7400 Kaposvár, Hungary; bene.szabolcs.albin@uni-mate.hu (S.A.B.); anton.istvan@uni-mate.hu (I.A.)

* Correspondence: zsolnai.attila@uni-mate.hu or attila.zsolnai@gmail.com

Simple Summary: By analyzing the genome of Hungarian Holstein Friesian cows, we looked for genomic regions which have an effect on the milk, fat, and protein yield. Among the sampled animals and the investigated nucleotides, nine were simultaneously associated with milk, fat, and protein yield. Among the nine variants, two had opposite effects; for example, while increasing the value of milk yield, the other one or two parameters' values decreased. The acquired knowledge can help in the planning of breeding schemes to avoid unwanted interactions among the abovementioned yield parameters.

Abstract: This study aimed to find SNPs that have an effect on the estimated breeding values (EBVs) of milk (MY), fat (FY), and protein yield (PY) of Holstein Friesian cows in Hungary. Holstein Friesian cows (n = 2963) were genotyped on a Eurogenomics (EuroG_MDv4) chip. The EBVs for MY, FY, and PY were obtained from the Association of Hungarian Holstein Breeders (AHHB). The loci associated with the EBVs were identified via three approaches: the calculation of genetic distance of the SNPs (F_{st_marker}), linear regression, and haplotype association tests. Nine SNPs were significantly associated with MY, FY, and PY located on BTA 2, 5, 28, and X. Among the nine SNPs identified, BTB-00219372 on BTA 5 had a positive β coefficient for MY and a negative β coefficient for FY and PY. In addition, BovineHD3000027615 on BTA X had a positive β coefficient for both MY and PY, as well as a negative β coefficient for FY. The identified SNPs were located near several genes that remain unstudied in cattle, which are potential targets for closer scrutiny in relation to milk properties. The markers associated with two or three EBVs could be used in selection with high efficiency to accelerate genetic development and help AHHB experts achieve their breeding. Most marker effects point in the same direction on EBVs; however, we found that BTB-00219372 and BovineHD3000027615 could be used with caution to increase one EBV while decreasing the other EBV or EBVs.

Keywords: Holstein; genome-wide association; single-nucleotide polymorphism; estimated breeding value



Citation: Bognár, L.; Kőrösi, Z.J.; Bene, S.A.; Szabó, F.; Anton, I.; Zsolnai, A. Simultaneous Effects of Single-Nucleotide Polymorphisms on the Estimated Breeding Value of Milk, Fat, and Protein Yield of Holstein Friesian Cows in Hungary. *Animals* **2024**, *14*, 3518. <https://doi.org/10.3390/ani14233518>

Academic Editor: Maria Luisa Dettori

Received: 3 October 2024

Revised: 2 December 2024

Accepted: 2 December 2024

Published: 5 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Holstein Friesian (HF) is perhaps the most recognized and widely distributed dairy cattle breed worldwide, originating in the Netherlands, where black Batavian and white Friesian animals were crossed to create a new breed with superior milk-producing abilities, even under limited feed resources [1]. The Holstein breed rapidly spread to North America in the 1800s and later to the entire world [1].

In 1960, The World Holstein Friesian Federation was founded to improve, develop, and promote the HF breed, with the first international Holstein meeting held in 1964 [2]. The HF breed was introduced to Hungary in the 1970s, and the Association of Hungarian Holstein Breeders (AHHB) was founded in 1989. At present, the AHHB has more than 900 members possessing over 248,000 cows [3].

Advances in molecular genetics, especially in the typing of single-nucleotide polymorphisms (SNPs) using the microarray technique, and the availability of genomic data on many individuals have radically changed the entire dairy sector worldwide [4]. DNA microarray-based genomic investigations on Holsteins in the United States started in 2009 and were quickly recognized as a valuable tool for selective breeding [5].

In the last few years, several genome-wide association studies (GWASs) were performed to identify the loci associated with different production traits in HF cattle. For instance, in Chinese HF, several SNPs were linked to various milk production traits [6], milk fatty acid [7], milk protein composition and/or protein percentage [8], udder health, and conformation [9,10]. In Irish Holsteins, novel SNPs have been associated with milk production using advanced statistical methods [11]. In Nordic Holsteins, SNPs were revealed [12] affecting both milk production and mastitis resistance. A large-scale GWAS in US Holsteins has confirmed previously identified SNPs and uncovered new genetic effects on various milk production traits [13]. Another large-scale study in European countries identified regions on the *Bos taurus* genome associated with milk yield (MY) and lactation curve parameters, supporting both known and novel candidate genes for MY in HF cows [14].

In Hungary, Anton et al. [15] investigated the effect of the lysine/alanine (K232A) polymorphism at the *DGAT1* locus on milk production traits of HF cows. The first GWASs related to cattle production in Hungary were published in 2018, identifying several loci associated with the breeding values for fertility and beef [16]. A major breakthrough in Hungarian HF came with the introduction of genomically enhanced breeding value (GEBV) estimation in 2017, which blended the genomic best linear unbiased prediction with traditional breeding values [17], followed by the HUNGENOM project in 2019, providing a genomic selection tool for Hungarian breeders based on genomic breeding value estimation [3].

This study, based on the data collected in the HUNGENOM project, aimed to reveal the combined effects of SNPs on the previously calculated EBVs for milk (MY), fat (FY), and protein yield (PY) in Hungarian HF cows.

2. Materials and Methods

This study did not require approval from the Ethical Committee on Animal Experiments, since genotyping is part of the routine breeding procedure of the AHHB. All data concerning the phenotype and genotype of animals were provided by the AHHB. HF cows were genotyped using the EuroG_MDv4 microarray chip (Eurogenomics, Amsterdam, The Netherlands) containing 67,227 SNPs. After including only samples with a call rate greater than 0.95 and SNPs with a call rate above 0.95, as well as removing duplicated samples, the final dataset comprised 2963 cows and 59,151 SNPs.

EBVs were calculated by the AHHB using 40,947 SNPs based on the genomic breeding value estimation method specified for Hungarian data [18], excluding SNPs located on chromosome X. The genomic prediction model was based on the Bayesian multi-QTL model [19], where the effects of dense SNPs across the whole genome were fitted directly without using haplotypes or identical-by-descent probabilities [20]. Although the method can be applied for multiple traits simultaneously, the routine genomic evaluations are single-trait analyses, i.e., $m = 1$. For m traits, the following model was applied:

$$y_i = \mu + u_i + \sum_{j=1}^{40947} z_{ij} q_j v_j + e_i \quad (1)$$

where y_i is the $(m \times 1)$ vector of phenotypes (deregressed proofs), u_i is the $(m \times 1)$ vector of random polygenic effects, and e_i is the $(m \times 1)$ vector of residuals of animal i ; μ is the $(m \times 1)$ vector of fixed trait means; q_j is the (3×1) vector of random non-scaled SNP effects

for SNP j ([40947]) with alleles 0 (missing), 1 (A), and 2 (B); v_j is the $(1 \times m)$ random scaling vector for SNP j ; and z_{ij} is the (1×3) design vector for animal i and SNP j ($[0\ 2\ 0]$, $[0\ 1\ 1]$, $[0\ 0\ 2]$, or $[2\ 0\ 0]$ for homozygous [AA], heterozygous [AB], homozygous [BB], or non-genotyped [00] animals at SNP j , respectively).

For the GWASs, the animals were divided according to their EBVs for MY, FY, and PY (EBV_{milk} , EBV_{fat} , EBV_{prot} , respectively). First, in each EBV category, high and low valued groups were created as follows: $EBV_{milk_high} > 1465$, $EBV_{milk_low} < 328$; $EBV_{fat_high} > 65$, $EBV_{fat_low} < 19$; $EBV_{prot_high} > 51$, $EBV_{prot_low} < 21$. For each trait (EBV_{milk} , EBV_{fat} , EBV_{prot}), three independent approaches (genetic distance of the SNPs (Fst_marker), linear regression [21], and haplotype association tests [22]) were calculated using the SNP and Variation Suite 8.8.1 (SVS) software (Golden Helix, Bozhan, MT, USA). The flow diagram of the procedure is illustrated in Figure 1.

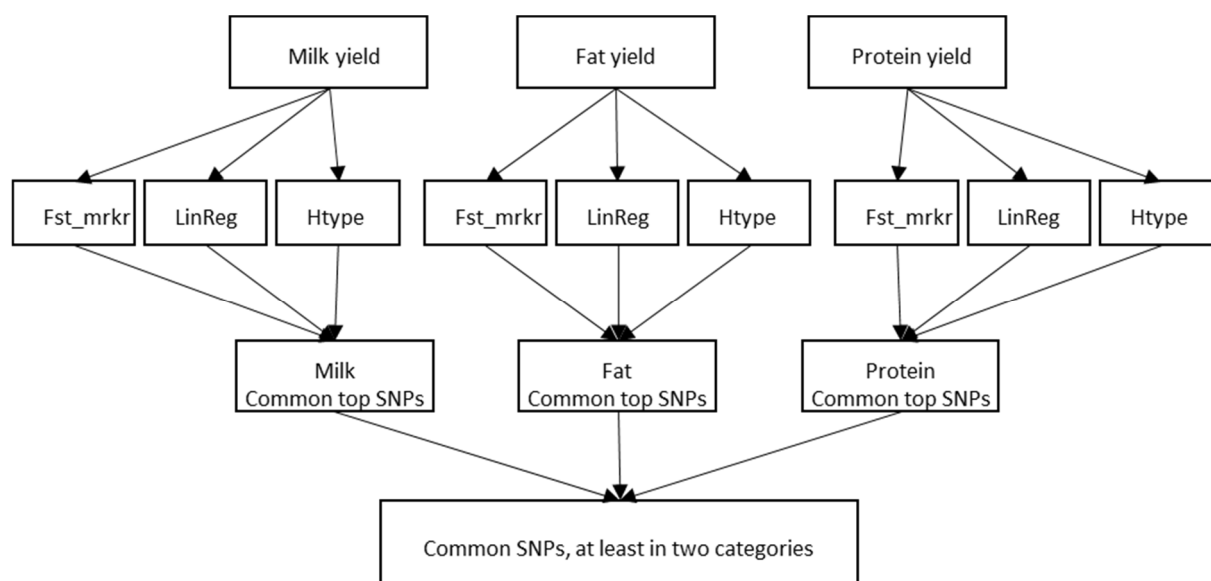


Figure 1. Flow diagram of the search for single-nucleotide polymorphisms (SNPs) associated with two or three of the estimated breeding values (EBVs). First row: EBV values for MY, FY, and PY obtained from the Association of Hungarian Holstein Breeders; second row: the applied tests for each trait (Fst_mrkr: genetic distance of SNPs, LinReg: linear regression, Htype: haplotype association); third row: the top SNPs found using different tests were identified for each trait; fourth row: the top SNPs shared by two or three traits were identified.

In haplotype association tests, the window size was set to 5 markers; a chi-squared test was performed for each haplotype. The haplotypes were constructed using the expectation maximization (EM) algorithm (maximum EM iteration = 50, EM convergence tolerance = 0.0001). After visual inspection of the Manhattan plots (Figure 2) of the EBV_{milk} , EBV_{fat} , and EBV_{prot} associations, the threshold values for F_{st_marker} and $-\log_{10}(p)$ in linear regression or haplotype association were 0.06, 8, and 8 for EBV_{milk} ; 0.06, 9, and 9 for EBV_{fat} ; and 0.08, 9, and 9 for EBV_{prot} , respectively. The SNPs above the thresholds were determined for each trait (third row in Figure 1), and those associated with at least two traits were identified (fourth row in Figure 1). The false discovery rates [23] of the identified 74 SNPs ranged from 1.3×10^{-21} to 6.0×10^{-06} . Indices were created for the F_{st_marker} , linear regression, and haplotype association analysis results by rescaling the values from 0 to 1 and averaging them.

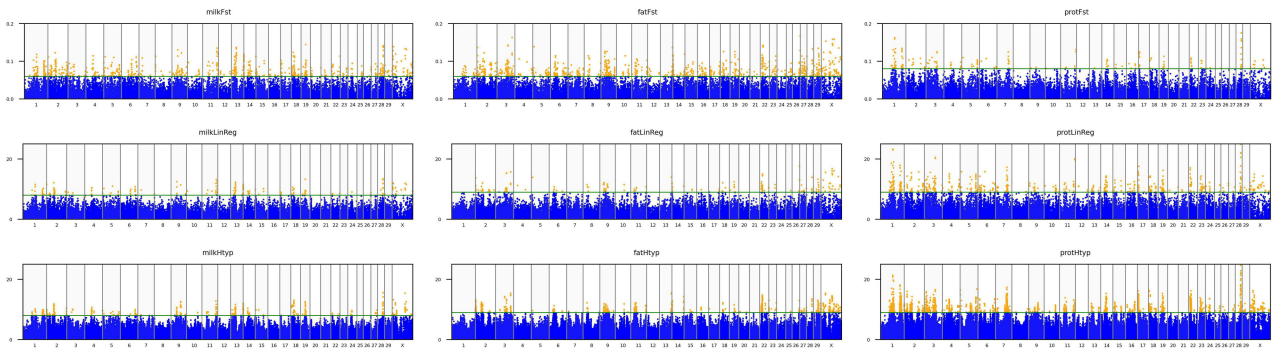


Figure 2. Manhattan plots of the associations of EBVs for milk (EBV_{milk} ; **left**), fat (EBV_{fat} ; **middle**), and protein (ENV_{prot} ; **right**) with F_{st_marker} (**top row**), linear regression (**middle row**), and five-SNP haplotypes (**bottom row**). The green lines are the thresholds above which the markers are considered as top hits.

The *B. taurus* genome assembly ARS-UCD1.2 was used to look for genes located ± 1 million base pairs (Mbp) from the common hits (Tables 1 and S1). When no gene was mapped within ± 1 Mbp, the distance was extended to ± 3 Mbp.

Table 1. The 74 SNPs associated with at least two of the examined traits.

SNP no.	EBV_{milk}	EBV_{fat}	EBV_{prot}
5	+	+	
44	+		+
16		+	+
9	+	+	+
Total	58	30	69

3. Results

Five SNPs associated with EBVs for MY and FY were identified on BTAs 9, 18, and 19. In addition, we discovered 44 SNPs associated with EBVs for MY and PY located on BTAs 1–6, 11, 13–15, 18, 19, 24, 28, and X. Moreover, 16 SNPs were related to EBVs for FY and PY on BTAs 3, 11, 19, 22, and X. Furthermore, nine SNPs were associated with EBVs for MY, FY, and PY located on BTAs 2, 5, 28, and X (Table 1, Supplementary Table S1). The maximum values of the identified SNPs were 0.17 for F_{st_marker} , 24.9 for the $-\log_{10}(p)$ of the linear regression, and 26.4 for the $-\log_{10}(p)$ of the haplotype association. Among our findings, the most prominent hits, mean values >0.8 (Supplementary Table S2), were located on BTAs 2, 11, 19, 28, and X for MY; BTAs 3, 22, 28, and X for FY; and BTAs 1 and 28 for PY. Among the nine SNPs associated with the EBVs for MY, FY, and PY, seven were among the top (means >0.8), three of them within 1.18 million base pairs on BTA 28.

3.1. SNPs Associated with Two EBVs and Their Surrounding Genes

The genes found in the vicinity of the top SNPs examined are summarized in Supplementary Table S1; the descriptions and references of them are given in Supplement File S1.

3.2. SNPs Associated with EBV_{milk} , EBV_{fat} , and EBV_{prot} and Their Surrounding Genes

Nine SNPs on BTAs 2, 5, 28, and X were associated with EBV_{milk} , EBV_{fat} , and EBV_{prot} . No genes were within ± 3 Mbp of the two SNPs on BTA 2 (Supplementary Table S1). The one SNP on BTA 5 was near two genes: *PPFIA2* and *METTL25*. On BTA 28, we identified three genes (*CCSER2*, *SHLD2*, and *ANXA8L1*), while on chromosome X, the abundance of the genes was elevated.

4. Discussion

This study aimed to identify the SNPs associated with ENVs for two or three milk traits. It used three algorithms to identify the associated SNPs (Figure 1) and used the top hits (Figure 2) to identify the candidate genes. Many of the surrounding genes have already been studied in *B. taurus*, *B. indicus*, or *Bos grunniens* and are associated with milk characteristics, weight, collagen synthesis, or sperm quality.

On BTA 5, *PPFIA2* has been associated with regulation of the reproduction process [24]. The gene was identified as a candidate gene for 305-day MY in Guzerá cattle [25]. Furthermore, it was also associated with MY and FY in several Thai dairy cattle populations [26]. *METTL25* was identified as a candidate gene for claw disorder digital dermatitis in HF and Simmental cows and may affect disease resistance [27].

On BTA 28, mutations in *CCSER2* were found to affect the fat, protein, casein, and lactose traits of Gannan yak milk [28]. *SHLD2* is an effector of transformation-related protein 53 binding protein 1 (TRP53BP1/53BP1) and was critical in suppressing large deletions within the immunoglobulin heavy-chain locus in mice [29]. The gene ontology annotations associated with *ANXA8L1* included calcium ion binding and calcium-dependent phospholipid binding [30].

On BTA X, *FMR1* has been associated with bull fertility traits [31] and fragile X syndrome [32]. *FMR1NB* has an unknown function that is predicted to be a membrane protein [33]. *AFF2* is implicated in fragile X syndrome in Nelore cattle [34]. Mutations in *IDS* cause mucopolysaccharidosis type II, also known as the Hunter syndrome [35]. *AKAP4* is expressed in various tissues and may play a role in defects in sperm flagellum and motility [36]. *CCNB3* is expressed in various tissues and is indispensable for female fertility in mice [37]. *DGKK* was associated with hypospadias in humans and Holstein cattle, which is a congenital defect of the genital region [38]. SNPs in *SHROOM4* have been associated with intellectual developmental disorder and epilepsy [39]. Mutations in *BMP15* have been associated with fresh sperm motility in Holstein bulls, making it a potential marker for sperm quality [40]. *NUDT10* and *NUDT11* are luteinizing hormone-regulated genes in bovine granulosa and have major roles in ovarian function in Holstein cows [41]. *CXHXorf67* is associated with endometrial stromal sarcomas [42]. *GSPT2* is involved in translation termination and mRNA decay. It may be involved in mRNA stability [43]. *MAGED1* and *MAGED4B* showed very high expression during estrus [44]. Mutations in *HEPH* can cause severe microcytic anemia in mice [45]. *VSIG4* was associated with macrophage activation by regulating the pyruvate metabolism of mitochondria [46]. *MSN* has roles in lymphocyte homeostasis and primary immunodeficiency diseases [47]. *LAS1L* may be involved in neurogenetic disorders in humans [48]. *ZC3H12B* is involved in the proinflammatory activation of macrophages [49]. *ZC4H2* has been identified as a candidate gene for semen quality and fertility in Egyptian buffalo bulls [50]. *ASB12* may play a role in muscle fiber growth in different cattle breeds [51]. *AMER1* is a potential candidate gene for X-linked hereditary diseases in cattle [52]. *ARHGEF9* is a potential candidate gene for X-linked hereditary diseases and cognitive impairment in cattle [52]. In humans, *SPIN4* is associated with overgrowth syndrome and hyperekplexia [53]. Markers covering the *ZXDB* region were highly differentiated in German Mutton compared to Dorper and Sunit sheep [54]. *ZNF674* is implicated in non-syndromic X-linked cognitive disabilities in humans [55].

In our study, *DGAT1* was not among the top hits due to our filtering setup of the association results. As shown in Figure 2, the applied threshold on the Manhattan plots for PY did not allow the peak at the beginning of BTA 14 to be included in our comparison. Regarding FY, of the algorithms and filters applied, only haplotype regression identified the *DGAT1* region. However, this region was retained among the top hits for MY. On BTA 2, two SNPs associated with the EBVs for MY, FY, and PY (Hapmap47966-BTA-47563, ARS-BFGL-NGS-113042) were 2 Mbp from markers associated with MY reported by Minozzi [56].

Atashi et al. [14] identified three regions on *Bos taurus* autosome (BTA) 14 containing SNPs associated with MY. Our combined hits associated with EBVs for at least two of the

examined traits were not within 3 Mbp of these genes. Jiang et al. [6] identified several SNPs associated with multiple milk production traits in Chinese Holsteins, where three SNPs on BTA14 were associated with MY, FY, and PY. Meredith et al. [11] identified no significant SNPs associated with MY, FY, and PY in Irish HF cows. In the Hungarian Holstein population, we did not find SNPs on BTA14 associated with these traits simultaneously.

Jiang et al. [13] reported three genes (solute carrier family 4 member 4 [SLC4A4], ADAM metalloproteinase with thrombospondin type 1 motif 3 [ADAMTS3], and GC vitamin D binding protein [GC]) on BTA6 that had significant additive effects on the MY and PY of US Holstein cattle. In addition, two SNPs on BTA5 (rs41257416 [position: 105,870,613] and rs110000229 [position: 105,804,923]), located very close to our SNPs, had significant additive effects on PY. Pedrosa et al. [57] described 98 genes located on BTA14 associated with milk production traits (MY, FY, PY, FP, and PP) in North American Holstein cattle. A review article published by Bekele et al. [58] mentioned 136 SNPs significantly associated with two or more milk production traits in Holstein cattle and crossbreds. Out of them, fifty-three, eighteen, ten, and seven SNPs were located on BTAs 14, 6, 20, and 1, respectively. Our top SNPs associated with three traits were located on BTAs 2, 5, 28, and X.

Kolenda et al. [59] found that the *PAEP* gene (beta-lactoglobulin) was associated with MY, FY, and PY. In our study, we identified a marker close to this gene, which was associated with EBV_{fat} and EBV_{prot}. Regarding EBV_{milk}, the closest marker on BTA11 was 5.5 Mbp from *PAEP* (Supplementary Table S1).

PPFIA2 has been identified as a candidate gene in connection with MY and FY [25,26], and it takes part in the regulation of reproduction [24]. As for *PPFIA2*, we found a connection not just with EBV_{fat} and EBV_{prot} but with EBV_{milk} as well.

Several genes around the top hits either had not been studied in cattle and/or had an unknown function. In other cases, their described function in other mammals suggests influences on EBVs. We believe all genes near the associated SNPs warrant more detailed functional studies. In Supplementary Table S1, we underlined those candidate genes that take part in calcium transport (*EFCAB10*, *SBSPON*, *JPH1*, *EMC8*, *CACNA1F*, *ANXA8L1*), starred genes with collagen-related processes (*TAPBPL**, *P4HA3**), and cfm-labeled some genes known to have cilia and flagella functions (*SPACA9^{cfm}*, *B9D^{cfm}*, *CCDC13^{cfm}*) and/or known to be membrane proteins.

We were also interested in whether the regression coefficients of the SNPs reported here were all positive or negative across the three EBVs. While most regression coefficients were found to be in the same direction (i.e., consistently negative or positive for EBV_{milk}, EBV_{fat}, and EBV_{prot}), two SNPs showed opposite signs. BTB-00219372 on BTA 5 had a positive β coefficient for MY but a negative β coefficient for FY and PY. In addition, BovineHD3000027615 on BTA X had a positive β coefficient for MY and PY but a negative β coefficient for FY (Supplementary Table S2). The first case might reflect cases where cows produce more but slim milk, which can be a problem in profit realization, including calf rearing. The second case means higher MY and PY but lower FY, which might affect the enjoyment value of milk [60] and the quality of the cheese products.

Since not just linear regression was used to determine the common sets of SNPs, we standardized the results of the three algorithms by rescaling their values between 0 and 1 and averaging them for each EBV. The means of the rescaled values for the top alleles associated with EBVs for at least two of the studied traits ranged from 0.445 to 0.9260. The strongest signal was for SNP BTA-64158-no-rs on BTA 28 (Supplementary Table S2, Figure 2) with ENVs for all three examined traits. Supplementary Table S2 and Figure 3 show that 58 SNPs were significantly associated with EBV_{milk} (orange bars), while 16 SNPs were significantly associated with EBV_{fat} (blue bars). We identified 5 SNPs associated with the EBVs for MY and FY (BTAs 9, 18, 19), 44 SNPs associated with the EBVs for MY and PY (BTAs 1–6, 11, 13–15, 18, 19, 24, 28, and X), and 16 SNPs associated with the EBVs for FY and PY (BTAs 3, 11, 19, 22, and X).

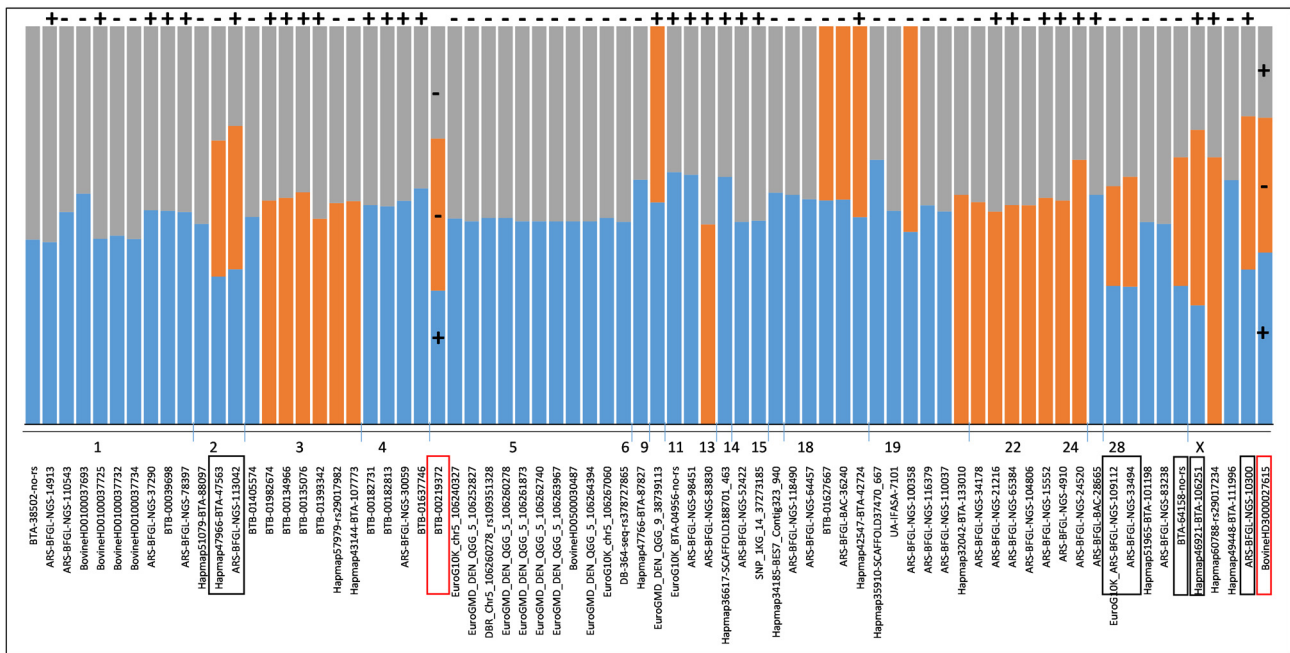


Figure 3. The 74 SNPs associated with EBVs for two or three of the examined traits. The orange (EBV_{milk}), blue (EBV_{fat}), and gray (EBV_{prot}) bars correspond to the regression β coefficients of each SNP (Supplementary Table S2). SNPs associated with all three EBVs are shown in boxes. At the top of the figure, the plus and the minus denote all positive or all negative β coefficients across traits. Red boxes denote SNPs where the β coefficients were opposite, e.g., FY and PY were negative while MY was positive on BTA 5, and FY was negative while MY and PY were positive on chromosome X.

5. Conclusions

Our study exclusively focused on the SNPs and candidate genes associated with EBVs for two or three milk production traits (MY, PY, and FY) in HF cows in Hungary. Notably, nine SNPs were associated with the EBVs for MY, FY, and PY (BTAs 2, 5, 28, and X). Some of the identified markers were located very near to previously reported chromosome regions or genes that were not previously linked to milk parameters in cattle or linked to other properties. Other genes reported to have an effect on and/or that could be linked to milk properties (*SBSPON*, *KLHL8*, *SLC35A2*, *SLC38A5*, *CTH*, *SPACA9*, *PAEP*, *CTNNB1*, *OXTR*, *PIN4*, *PPFIA2*, and *CCSER2*) were also found in our study (for genes *KLHL8*, *SLC35A2*, *SLC38A5*, *CTH*, *CTNNB1*, *OXTR*, and *PIN4*, see Supplementary Table S1 and Supplement File S1). Several genes like *EFCAB10*, *GLOD5*, *NONO*, and *TMEM70* were reported in this study and have not been previously investigated in connection with milk properties in cattle.

Regarding the marker effects, they exhibited a consistent directional influence across EBVs. However, two markers, BTB-00219372 and BovineHD3000027615, showed divergent effects. Since these markers could be used to selectively increase one EBV while potentially decreasing others, special consideration when using these markers in selective breeding programs is required.

These findings provide new information that could accelerate genetic progress and may help specialists from the AHHB achieve their breeding and selection goals in the Holstein population in Hungary.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani14233518/s1>, Table S1: The names of the markers associated with two or three EBV values, their genomic positions (B. taurus genome build ARS-UCD1.2), and genes found around the markers; Table S2: The means of rescaled values of the applied algorithms and regression beta values of linear regressions on EBVs for MY, FY, and PY; Supplement File S1: SNPs Associated with Two EBVs and Their Surrounding Genes.

Author Contributions: L.B., I.A. and A.Z. were responsible for the conceptualization and design of the study; L.B., Z.J.K., S.A.B. and F.S. performed data acquisition; L.B., I.A. and A.Z. analyzed and/or interpreted the data; L.B., I.A. and A.Z. drafted the manuscript; L.B., Z.J.K., S.A.B., F.S., I.A. and A.Z. reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study did not require approval from the Ethical Committee on Animal Experiments, since genotyping is part of the routine breeding procedure of the AHHB. The study included 2963 cows. All data concerning the phenotype and genotype of animals were provided by the AHHB.

Informed Consent Statement: No additional informed consent forms were required from individual participants or animal owners in this study. The samples were acquired from the National Breeding Program managed by the National Association of Hungarian Holstein Friesian Breeders. All members who joined the association agreed to the terms, which included providing the necessary data to ensure the success of the breeding program.

Data Availability Statement: The raw dataset presented in this article is not readily available because the data are part of ongoing studies and are owned by the Association of Hungarian Holstein Breeders. Requests to access the datasets should be directed to the first author.

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Elischer, M. *History of Dairy Cow Breeds: Holstein*; Michigan State University Extension: Howell, MI, USA, 2014.
2. World Holstein Friesian Federation. Available online: <https://whff.info/conferences/> (accessed on 21 July 2024).
3. Association of Hungarian Holstein Breeders. Available online: <https://www.holstein.hu> (accessed on 21 July 2024).
4. Nayeri, S.; Sargolzaei, M.; Abo-Ismael, M.K.; May, N.; Miller, S.P.; Schenkel, F.; Moore, S.S.; Stothard, P. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet.* **2016**, *17*, 75. [[CrossRef](#)] [[PubMed](#)]
5. Wiggans, G.R.; Carrillo, J.A. Genomic selection in United States dairy cattle. *Front. Genet.* **2022**, *13*, 994466. [[CrossRef](#)] [[PubMed](#)]
6. Jiang, L.; Liu, J.; Sun, D.; Ma, P.; Ding, X.; Yu, Y.; Zhang, Q. Genome wide association studies for milk production traits in Chinese Holstein population. *PLoS ONE* **2010**, *5*, e13661. [[CrossRef](#)] [[PubMed](#)]
7. Li, C.; Sun, D.; Zhang, S.; Wang, S.; Wu, X.; Zhang, Q.; Liu, L.; Li, Y.; Qiao, L. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. *PLoS ONE* **2014**, *9*, e96186. [[CrossRef](#)]
8. Zhou, C.; Li, C.; Cai, W.; Liu, S.; Yin, H.; Shi, S.; Zhang, Q.; Zhang, S. Genome-Wide Association Study for Milk Protein Composition Traits in a Chinese Holstein Population Using a Single-Step Approach. *Front. Genet.* **2019**, *10*, 72. [[CrossRef](#)]
9. Wang, X.; Ma, P.; Liu, J.; Zhang, Q.; Zhang, Y.; Ding, X.; Jiang, L.; Wang, Y.; Zhang, Y.; Sun, D.; et al. Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC Genet.* **2015**, *16*, 111. [[CrossRef](#)]
10. Nazar, M.; Abdalla, I.M.; Chen, Z.; Ullah, N.; Liang, Y.; Chu, S.; Xu, T.; Mao, Y.; Yang, Z.; Lu, X. Genome-Wide Association Study for Udder Conformation Traits in Chinese Holstein Cattle. *Animals* **2022**, *12*, 2542. [[CrossRef](#)]
11. Meredith, B.K.; Kearney, F.J.; Finlay, E.K.; Bradley, D.G.; Fahey, A.G.; Berry, D.P.; Lynn, D.J. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genet.* **2012**, *13*, 21. [[CrossRef](#)]
12. Cai, Z.; Dusza, M.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. *Genet. Sel. Evol.* **2020**, *52*, 19. [[CrossRef](#)]
13. Jiang, J.; Ma, L.; Prakapenka, D.; VanRaden, P.M.; Cole, J.B.; Da, Y. A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle. *Front. Genet.* **2019**, *10*, 412. [[CrossRef](#)]
14. Atashi, H.; Salavati, M.; De Koster, J.; Ehrlich, J.; Crowe, M.; Opsomer, G.; Gplus, E.c.; Hostens, M. Genome-wide association for milk production and lactation curve parameters in Holstein dairy cows. *J. Anim. Breed. Genet.* **2020**, *137*, 292–304. [[CrossRef](#)] [[PubMed](#)]
15. Anton, I.; Kovacs, K.; Fesus, L.; Varhegyi, J.; Lehel, L.; Hajda, Z.; Polgar, J.P.; Szabo, F.; Zsolnai, A. Effect of DGAT1 and TG gene polymorphisms on intramuscular fat and on milk production traits in different cattle breeds in Hungary. *Acta Vet. Hung.* **2008**, *56*, 181–186. [[CrossRef](#)] [[PubMed](#)]
16. Anton, I.; Huth, B.; Fuller, I.; Gabor, G.; Hollo, G.; Zsolnai, A. Effect of single-nucleotide polymorphisms on the breeding value of fertility and breeding value of beef in Hungarian Simmental cattle. *Acta Vet. Hung.* **2018**, *66*, 215–225. [[CrossRef](#)] [[PubMed](#)]
17. Bengtsson, C.; Stalhammar, H.; Strandberg, E.; Eriksson, S.; Fikse, W.F. Association of genomically enhanced and parent average breeding values with cow performance in Nordic dairy cattle. *J. Dairy. Sci.* **2020**, *103*, 6383–6391. [[CrossRef](#)]
18. Stoop, W.M.; Eding, H.; Schrooten, C. Method of Genomic Breeding Value Estimation, Specified for Hungarian data. In *Proposal for the National Association of Hungarian Holstein Friesian Breeders*; Working Document; CRV BV: Arnhem, The Netherlands, 2017.

19. Meuwissen, T.H.; Goddard, M.E. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genet. Sel. Evol.* **2004**, *36*, 261–279. [[CrossRef](#)]
20. Calus, M.P.; Meuwissen, T.H.; de Roos, A.P.; Veerkamp, R.F. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* **2008**, *178*, 553–561. [[CrossRef](#)]
21. Vilhjalmsón, B.J. Mixmogam [Internet]. Available online: <https://github.com/bvilhjal/mixmogam> (accessed on 21 July 2024).
22. Excoffier, L.; Slatkin, M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* **1995**, *12*, 921–927. [[CrossRef](#)]
23. Storey, J.D. A direct approach to false discovery rates. *J. R. Stat. Soc. B* **2002**, *64*, 479–498. [[CrossRef](#)]
24. Mota, L.F.M.; Lopes, F.B.; Fernandes Junior, G.A.; Rosa, G.J.M.; Magalhaes, A.F.B.; Carvalheiro, R.; Albuquerque, L.G. Genome-wide scan highlights the role of candidate genes on phenotypic plasticity for age at first calving in Nelore heifers. *Sci. Rep.* **2020**, *10*, 6481. [[CrossRef](#)]
25. Paiva, J.T.; Peixoto, M.G.C.D.; Bruneli, F.A.T.; Alvarenga, A.B.; Oliveira, H.R.; Silva, A.A.; Silva, D.A.; Veroneze, R.; Silva, F.F.; Lopes, P.S. Genetic parameters, genome-wide association and gene networks for milk and reproductive traits in Guzerá cattle. *Livest. Sci.* **2020**, *242*, 104273. [[CrossRef](#)]
26. Laodim, T.; Koonawootrittriron, S.; Elzo, M.A.; Suwanasopee, T.; Jattawa, D.; Sarakul, M. Genetic factors influencing milk and fat yields in tropically adapted dairy cattle: Insights from quantitative trait loci analysis and gene associations. *Anim. Biosci.* **2024**, *37*, 576–590. [[CrossRef](#)] [[PubMed](#)]
27. Solzer, N.; Brügemann, K.; Yin, T.; König, S. Genetic evaluations and genome-wide association studies for specific digital dermatitis diagnoses in dairy cows considering genotype x housing system interactions. *J. Dairy. Sci.* **2024**, *107*, 3724–3737. [[CrossRef](#)] [[PubMed](#)]
28. Yang, G.; Zhang, J.; Ma, X.; Ma, R.; Shen, J.; Liu, M.; Yu, D.; Feng, F.; Huang, C.; Ma, X.; et al. Polymorphisms of CCSE1 Gene and Their Correlation with Milk Quality Traits in Gannan Yak (*Bos grunniens*). *Foods* **2023**, *12*, 4318. [[CrossRef](#)] [[PubMed](#)]
29. Ling, A.K.; Munro, M.; Chaudhary, N.; Li, C.; Berru, M.; Wu, B.; Durocher, D.; Martin, A. SHLD2 promotes class switch recombination by preventing inactivating deletions within the Igh locus. *EMBO Rep.* **2020**, *21*, e49823. [[CrossRef](#)]
30. ANXA8L1 GeneCards. Available online: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=ANXA8L1> (accessed on 21 July 2024).
31. Fonseca, P.A.S.; Dos Santos, F.C.; Lam, S.; Suarez-Vega, A.; Miglior, F.; Schenkel, F.S.; Diniz, L.A.F.; Id-Lahoucine, S.; Carvalho, M.R.S.; Canovas, A. Genetic mechanisms underlying spermatic and testicular traits within and among cattle breeds: Systematic review and prioritization of GWAS results. *J. Anim. Sci.* **2018**, *96*, 4978–4999. [[CrossRef](#)] [[PubMed](#)]
32. Salcedo-Arellano, M.J.; Hagerman, R.J.; Martinez-Cerdeno, V. Fragile X syndrome: Clinical presentation, pathology and treatment. *Gac. Med. Mex.* **2020**, *156*, 60–66. [[CrossRef](#)]
33. FMR1NB GeneCards. Available online: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=FMR1NB> (accessed on 21 July 2024).
34. Hermisdorff, I.D.C.; Diaz, I.; de Camargo, G.M.F.; de Albuquerque, L.G.; Costa, R.B. Effect of genomic X-chromosome regions on Nelore bull fertility. *J. Appl. Genet.* **2021**, *62*, 655–659. [[CrossRef](#)]
35. Friso, A.; Tomanin, R.; Zanetti, A.; Mennuni, C.; Calvaruso, F.; La Monica, N.; Marin, O.; Zacchello, F.; Scarpa, M. Gene therapy of Hunter syndrome: Evaluation of the efficiency of muscle electro gene transfer for the production and release of recombinant iduronate-2-sulfatase (IDS). *Biochim. Biophys. Acta* **2008**, *1782*, 574–580. [[CrossRef](#)]
36. Miki, K.; Willis, W.D.; Brown, P.R.; Goulding, E.H.; Fulcher, K.D.; Eddy, E.M. Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. *Dev. Biol.* **2002**, *248*, 331–342. [[CrossRef](#)]
37. Chotiner, J.Y.; Leu, N.A.; Xu, Y.; Wang, P.J. Recurrent pregnancy loss in mice lacking the X-linked Ccnb3 genedagger. *Biol. Reprod.* **2022**, *106*, 382–384. [[CrossRef](#)]
38. Iannuzzi, A.; Braun, M.; Genuardo, V.; Perucatti, A.; Reinartz, S.; Proios, I.; Heppelmann, M.; Rehage, J.; Hulskotter, K.; Beineke, A.; et al. Clinical, cytogenetic and molecular genetic characterization of a tandem fusion translocation in a male Holstein cattle with congenital hypospadias and a ventricular septal defect. *PLoS ONE* **2020**, *15*, e0227117. [[CrossRef](#)] [[PubMed](#)]
39. Bian, W.J.; Li, Z.J.; Wang, J.; Luo, S.; Li, B.M.; Gao, L.D.; He, N.; Yi, Y.H. SHROOM4 Variants Are Associated With X-Linked Epilepsy With Features of Generalized Seizures or Generalized Discharges. *Front. Mol. Neurosci.* **2022**, *15*, 862480. [[CrossRef](#)] [[PubMed](#)]
40. Sun, L.P.; Song, Y.P.; Du, Q.Z.; Song, L.W.; Tian, Y.Z.; Zhang, S.L.; Hua, G.H.; Yang, L.G. Polymorphisms in the bone morphogenetic protein 15 gene and their effect on sperm quality traits in Chinese Holstein bulls. *Genet. Mol. Res.* **2014**, *13*, 1805–1812. [[CrossRef](#)] [[PubMed](#)]
41. Christenson, L.K.; Gunewardena, S.; Hong, X.; Spitschak, M.; Baufeld, A.; Vanselow, J. Research resource: Preovulatory LH surge effects on follicular theca and granulosa transcriptomes. *Mol. Endocrinol.* **2013**, *27*, 1153–1171. [[CrossRef](#)] [[PubMed](#)]
42. Dewaele, B.; Przybyl, J.; Quattrone, A.; Finalet Ferreira, J.; Vanspauwen, V.; Geerdens, E.; Gianfelici, V.; Kalender, Z.; Wozniak, A.; Moerman, P.; et al. Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma. *Int. J. Cancer* **2014**, *134*, 1112–1122. [[CrossRef](#)]
43. Zhouravleva, G.; Schepachev, V.; Petrova, A.; Tarasov, O.; Inge-Vechtomov, S. Evolution of translation termination factor eRF3: Is GSPT2 generated by retrotransposition of GSPT1's mRNA? *IUBMB Life* **2006**, *58*, 199–202. [[CrossRef](#)]

44. Mitko, K.; Ulbrich, S.E.; Wenigerkind, H.; Sinowatz, F.; Blum, H.; Wolf, E.; Bauersachs, S. Dynamic changes in messenger RNA profiles of bovine endometrium during the oestrous cycle. *Reproduction* **2008**, *135*, 225–240. [[CrossRef](#)]
45. Fuqua, B.K.; Lu, Y.; Darshan, D.; Frazer, D.M.; Wilkins, S.J.; Wolkow, N.; Bell, A.G.; Hsu, J.; Yu, C.C.; Chen, H.; et al. The multicopper ferroxidase hephaestin enhances intestinal iron absorption in mice. *PLoS ONE* **2014**, *9*, e98792. [[CrossRef](#)]
46. Li, J.; Diao, B.; Guo, S.; Huang, X.; Yang, C.; Feng, Z.; Yan, W.; Ning, Q.; Zheng, L.; Chen, Y.; et al. VSIG4 inhibits proinflammatory macrophage activation by reprogramming mitochondrial pyruvate metabolism. *Nat. Commun.* **2017**, *8*, 1322. [[CrossRef](#)]
47. Kovacs, A.L.; Karteszi, J.; Prohaszka, Z.; Kalmar, T.; Kesmarky, G.; Koltai, K.; Nagy, Z.; Sebok, J.; Vas, T.; Molnar, K.; et al. Hemizygous nonsense variant in the moesin gene (MSN) leads to a new autoimmune phenotype of Immunodeficiency 50. *Front. Immunol.* **2022**, *13*, 919411. [[CrossRef](#)]
48. Aguilera, C.; Gabau, E.; Ramirez-Mallafre, A.; Brun-Gasca, C.; Dominguez-Carral, J.; Delgadillo, V.; Laurie, S.; Derdak, S.; Padilla, N.; de la Cruz, X.; et al. New genes involved in Angelman syndrome-like: Expanding the genetic spectrum. *PLoS ONE* **2021**, *16*, e0258766. [[CrossRef](#)] [[PubMed](#)]
49. Wawro, M.; Wawro, K.; Kochan, J.; Solecka, A.; Sowinska, W.; Lichawska-Cieslar, A.; Jura, J.; Kasza, A. ZC3H12B/MCPIP2, a new active member of the ZC3H12 family. *RNA* **2019**, *25*, 840–856. [[CrossRef](#)] [[PubMed](#)]
50. El Nagar, A.G.; Salem, M.M.I.; Amin, A.M.S.; Khalil, M.H.; Ashour, A.F.; Hegazy, M.M.; Abdel-Shafy, H. A Single-Step Genome-Wide Association Study for Semen Traits of Egyptian Buffalo Bulls. *Animals* **2023**, *13*, 3758. [[CrossRef](#)] [[PubMed](#)]
51. Bazile, J.; Jaffrezic, F.; Dehais, P.; Reichstadt, M.; Klopp, C.; Laloe, D.; Bonnet, M. Molecular signatures of muscle growth and composition deciphered by the meta-analysis of age-related public transcriptomics data. *Physiol. Genom.* **2020**, *52*, 322–332. [[CrossRef](#)] [[PubMed](#)]
52. Id-Lahoucine, S.; Casellas, J.; Fonseca, P.A.S.; Suarez-Vega, A.; Schenkel, F.S.; Canovas, A. Deviations from Mendelian Inheritance on Bovine X-Chromosome Revealing Recombination, Sex-of-Offspring Effects and Fertility-Related Candidate Genes. *Genes* **2022**, *13*, 2322. [[CrossRef](#)]
53. Lui, J.C.; Wagner, J.; Zhou, E.; Dong, L.; Barnes, K.M.; Jee, Y.H.; Baron, J. Loss-of-function variant in SPIN4 causes an X-linked overgrowth syndrome. *JCI Insight* **2023**, *8*, e167074. [[CrossRef](#)]
54. Zhu, C.; Fan, H.; Yuan, Z.; Hu, S.; Zhang, L.; Wei, C.; Zhang, Q.; Zhao, F.; Du, L. Detection of Selection Signatures on the X Chromosome in Three Sheep Breeds. *Int. J. Mol. Sci.* **2015**, *16*, 20360–20374. [[CrossRef](#)]
55. Lugtenberg, D.; Yntema, H.G.; Banning, M.J.; Oudakker, A.R.; Firth, H.V.; Willatt, L.; Raynaud, M.; Kleefstra, T.; Fryns, J.P.; Ropers, H.H.; et al. ZNF674: A new kruppel-associated box-containing zinc-finger gene involved in nonsyndromic X-linked mental retardation. *Am. J. Hum. Genet.* **2006**, *78*, 265–278. [[CrossRef](#)]
56. Minozzi, G.; Nicolazzi, E.L.; Stella, A.; Biffani, S.; Negrini, R.; Lazzari, B.; Ajmone-Marsan, P.; Williams, J.L. Genome wide analysis of fertility and production traits in Italian Holstein cattle. *PLoS ONE* **2013**, *8*, e80219. [[CrossRef](#)]
57. Pedrosa, V.B.; Schenkel, F.S.; Chen, S.Y.; Oliveira, H.R.; Casey, T.M.; Melka, M.G.; Brito, L.F. Genomewide Association Analyses of Lactation Persistency and Milk Production Traits in Holstein Cattle Based on Imputed Whole-Genome Sequence Data. *Genes* **2021**, *12*, 1830. [[CrossRef](#)]
58. Bekele, R.; Taye, M.; Abebe, G.; Meseret, S. Genomic Regions and Candidate Genes Associated with Milk Production Traits in Holstein and Its Crossbred Cattle: A Review. *Int. J. Genom.* **2023**, *2023*, 8497453. [[CrossRef](#)] [[PubMed](#)]
59. Kolenda, M.; Sitkowska, B.; Kamola, D.; Lambert, B.D. Composite genotypes of progesterone-associated endometrial protein gene and their association with composition and quality of dairy cattle milk. *Anim. Biosci.* **2021**, *34*, 1283–1289. [[CrossRef](#)] [[PubMed](#)]
60. Pietrzak-Fiecko, R.; Kamelska-Sadowska, A.M. The Comparison of Nutritional Value of Human Milk with Other Mammals' Milk. *Nutrients* **2020**, *12*, 1404. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Management of “Modern” Holstein Cows Focusing on Sustainability and Resilience – Review of Recent Achievements

László Bognár^a, Ferenc Szabó^{b,*}

^aNational Association of Hungarian Holstein Friesian Breeders, 1134 Budapest, Hungary

^bDepartment of Animal Sciences, Albert Kázmér Faculty, Széchenyi István University, 9200 Mosonmagyaróvár, Hungary
szabo.ferenc@sze.hu

Keeping Holstein cows requires a strategic approach that maximises milk production and promotes sustainability and resilience. As global demand for dairy products continues to grow, it is increasingly important to balance the development of the industry with environmental protection. Integrating genomics and targeted breeding strategies in dairy production can significantly improve sustainability and efficiency. Genomic selection uses deoxyribonucleic acid (DNA) markers and single nucleotide polymorphism (SNP) information to predict the genetic value of an animal, allowing breeders to select beneficial traits such as disease resistance, longevity, fertility, and feed efficiency. This not only increases productivity but also reduces environmental burden. Tailored breeding strategies can improve herd health and productivity and ultimately reduce the environmental footprint per unit of milk. Managing “modern” Holstein cows with a focus on sustainability and resilience means implementing practices that promote animal welfare, minimise environmental impacts, and ensure the long-term viability of the dairy farm. In the context of modern Holstein cows, the term modern refers to Holstein cows that have been selectively bred and managed using contemporary agricultural practices and scientific advancements. These practices include the application of genomics, precision breeding programs, advanced dairy cattle health care techniques, and updated herd management practices. Modern Holstein cows have been adapted and bred to meet the demands of today’s dairy industry, considering factors like economic milk production, disease resistance, and environmental sustainability. The outcome of integrating genomics-targeted breeding and feeding strategies in Holstein cow management aims to enhance sustainability, productivity, and animal welfare, which would result in a more efficient and environmentally responsible dairy industry.

1. Introduction

The Holstein Friesian cattle is the dominant dairy breed in the world, including in Hungary, in terms of population number, level of production, and contribution to global milk production. Besides the advantages of intensive milk production using Holsteins, many experts and non-primarily experts criticise the sector for emitting greenhouse gases (GHG) and significantly contributing to global warming (Naranjo et al., 2020;). Numerous scientific works suggest that global warming, as a facet of climate change, represents a great challenge from a sustainability point of view (Peterson and Mitloehner, 2021). Several studies have indicated that livestock production, including transport, processing, and consumption, has a relatively large impact on climate change (Milani et al., 2011). According to the European Environment Agency (EEA), the EU-27’s agricultural GHG emissions in 2021 were 378,430 kt CO₂ equivalent (eq.), 11 % of total emissions. Livestock accounts for 245,448 kt CO₂ eq. (64.85 %), of which cattle enteric fermentation is responsible for 155,937 kt CO₂ eq. (63.53 %) and manure management for 28,613 kt CO₂ eq. (11.65 %). In Hungary, the agricultural sector contributes 12.55 % of total GHG emissions (7,202 kt CO₂ eq.). Within this, livestock accounts for 3,506 kt CO₂ eq. (42.43 %), with cattle enteric fermentation constituting 1,966 kt CO₂ eq. (56.07 %) and manure management, adding another 591 kt CO₂ eq. (16.85 %). Extensive research and practical experience suggest that promoting sustainable milk production from this breed could play a crucial role in mitigating climate change. Therefore, it is essential to focus on the management of Holstein cows, ensuring their high-quality production remains sustainable for the

future. This review aims to explore the key elements of sustainable management for Holstein cows, highlighting areas like genetic selection, nutrition, feed management, animal health, welfare, manure handling, waste management, energy efficiency, and water conservation.

2. Key Elements of Sustainability in the Dairy Sector

While many publications address the sustainability of intensive dairy production, the majority focus primarily on its impact on climate change. Although this impact is undoubtedly crucial, sustainability encompasses numerous other factors, some of which may be limiting. A more holistic approach to sustainability, considering multiple factors simultaneously, is needed.

This review seeks to explore various management elements contributing to the sustainability and resilience of Holstein cow milk production. By collectively addressing these elements, the environmental footprint of Holstein-based dairy operations can be further reduced. The key sustainable practices for managing Holstein cows discussed in this study include:

Genetic Selection: to balance production traits with factors like health, fertility, and longevity. Breeding programs should aim to develop cows that are adaptable to different environments, have good immune systems, and exhibit natural behaviours.

Nutrition and Feed Management: implementing sustainable feeding practices, such as incorporating locally sourced feeds, reducing reliance on imported feeds, and optimising ration formulations to minimise waste and environmental impact.

Animal Health and Welfare: Implement comprehensive animal health management protocols, vaccination programs, regular veterinary care, and preventive measures against common diseases.

Manure and Waste Management: implement effective manure management strategies, anaerobic digestion systems, composting, and nutrient management plans.

Energy Efficiency and Renewable Energy: Reduce the environmental footprint of the dairy operation by implementing energy-efficient practices. Optimise energy use in barns, milking parlours, and other facilities through efficient lighting, ventilation, and equipment.

Water Conservation: Implement water conservation practices, efficient watering systems, leak detection and repair, and responsible water use.

Knowledge Sharing: In the dairy industry, effective knowledge sharing is crucial for optimising production, enhancing sustainability, and navigating the evolving challenges and opportunities of the sector.

2.1 Genetic Selection

For a long time, the objective of genetic selection in the Holstein breed was to increase milk yield, butterfat, and protein. However, unfavourable genetic relationships among traits of great relevance to the industry (milk yield and fertility or welfare) have deteriorated some economically important traits, which has consequently motivated the development of more efficient breeding strategies for the increased long-term sustainability of the dairy cattle industry (Cole and VanRaden, 2018). Later on, the production aim was supplemented by conformation traits, with some functional traits such as longevity and calving ease (Brito et al., 2021). Together with intensive selection, the development of intensive dairy systems has been developed by innovations and technological breakthroughs, among which conventional genetic selection played a major role over the past decades (Miglior et al., 2017). Yet, the strong focus of the dairy industry on ensuring food security through higher productivity raises concerns about other sustainability dimensions (Clay et al., 2020). This situation requires a new breeding strategy, the simultaneous selection of productivity and functional traits such as adaptation, welfare, and resilience. Despite the major signs of progress in productivity, the long-term success of the dairy industry depends on the adoption of more sustainable breeding goals and management practices, especially from an agroecological perspective (Bito et al., 2021). The long-term sustainability of the dairy cattle industry depends on the development of balanced breeding goals to simultaneously improve animal health and welfare, productive efficiency, environmental impact, food quality and safety while minimising the loss of genetic diversity.

Genetic selection for some of these breeding goals has already been implemented around the world (Cole and VanRaden, 2018). Genetic selection is now based on a modern IT solution, DNA analysis and genomic breeding value estimation that allows us to make selection decisions not only on production, milk components (fat, protein), and functional conformation but also management traits with low heritability health and welfare, heat tolerance, adaptation and emission-related traits. The young animals could be genotyped at an early age, shortly after birth, pulling hair samples with follicles or even before birth using in vitro Embryo Transfer and embryonic cell biopsy techniques. The sampled DNA is analysed, and SNP information is used for calculating the Genomically Enhanced Breeding Values. This information provides a reliable tool for predicting their future performances in the dairy herd and can be used as a herd-size optimising practice to select or cull the given animals for dairy production. The minus variants could be inseminated with beef bulls, i.e., “Beef on Dairy”

program, while others with higher genetic merits would stay in the farm to raise the genetic level of the whole herd, capable of higher milk production more efficiently, reducing the environmental impact of milk production per produced units of milk. The selection index is the Holstein Global Index, which, besides production traits (Fat and Protein), includes health and management traits such as Somatic Cell Count Score, Productive Life, Feet and Legs, and Calving Ability. It is estimated from the Breeding Value of the individual traits. This selection scheme could help to meet the requirements of sustainability and resilience. Holstein cows have been selectively bred for high milk production. However, focusing solely on milk yield can lead to challenges in terms of sustainability and resilience. It is important to balance production traits with factors like health, fertility, and longevity.

Breeding programs should aim to develop cows that are adaptable to different environments, have good immune systems, and exhibit natural behaviours. Genetic selection has also been a major driver in increased productivity, longevity, and efficiency of dairy cows, further reducing the environmental impact per unit of milk production (Pryce and Haile-Mariam, 2020).

HUNGENOM Project The National Association of Hungarian Holstein Friesian Breeders successfully runs a genomics-based selection scheme for participating Holstein Dairy Farms, the HUNGENOM project. It has 73 active herds, 52,975 DNA (hair)samples analysed since the start of this program in 2019 and published 51,165 Breeding Values that serve as a base for making breeding/culling or crossbreeding decisions.

2.2 Nutrition and Feed Management

Feeding dairy cows is an important factor from an economic and environmental point of view. Feeding strategies and methods are important tools for improving the efficiency of milk production, as well as the emission of greenhouse gases, in order to help sustainability and welfare. There are many promising opportunities for further reducing emissions and helping sustainability through feed and waste additives (Martin et al., 2017.) Nutrition of dairy animals has also allowed for a substantial improvement in production via the use of total mixed-rations (TMR) balanced for nutrient and energy requirements accounting for each animal's age and stage of lactation (National Research Council, 2001). More energy-dense or more digestible feedstuffs result in additional energy available to the animal and generate less CH₄ from fermentation (Knapp et al., 2014). An increase in the starch proportion of the diet, such as through an increase in concentrate levels, also results in a more rapid fermentation of these feedstuffs and decreased CH₄ production (Johnson and Johnson, 1995). Feeding higher starch diets requires increased grain production, which can cause additional consumption of fossil fuel and fertilisers that results in an increase in N₂O and CO₂ (Johnson et al., 2002); however, this system is usually offset by the substantial decrease in overall in CH₄ emissions (Lovett et al., 2006). Feeding of cereal forages can also favour propionate production and reduce CH₄ emissions due to the higher starch concentration (Beauchemin et al., 2009). Higher concentrations of legumes, such as alfalfa, when compared with grass forage-based diets, can also lead to an overall decrease in CH₄ emissions (McCaughy et al., 1999). Age of harvest of forage also has a significant impact on emissions, with advancing maturity resulting in more lignified and less fermentable substrate contributing to increasing emissions associated with higher ruminal acetate (Pinares-Patiño et al., 2003). In addition to alterations in forage or concentrate composition and ratio, supplementation of lipids to dairy cattle diets can also mitigate enteric emissions (Hristov et al., 2013 a). Replacing concentrates with lipids results in a decrease in fermentable substrate by the microbes in the rumen and can also decrease total protozoa and methanogen populations (Ivan et al., 2004). An inclusion of high-oil by-products, such as distillers' grains or oilseed meals, can result in decreased CH₄ emissions (Hristov et al., 2013b). Research on ensiled feeds in relation to enteric emissions is generally lacking, although it is anticipated that corn silage will mitigate emissions due to its higher starch content (Gerber et al., 2013). When directly comparing grass-versus-corn silage, a higher inclusion of corn silage seems to mitigate enteric CH₄ emissions (Doreau et al., 2012). There are many potential methods to mitigate enteric emissions through alterations to nutrition strategy and composition (Hristov et al., 2015). Optimal nutrition plays a critical role not only in GHG emissions but also in the health and productivity of Holstein cows.

2.3 Animal Health and Welfare

Consumers are interested in the safety and quality of dairy products (Drake, 2007). However, more recently, there has been increased interest in the care and housing of cows that produce milk and its associated products (von Keyserlingk et al., 2009). A fundamental condition for the production of large-quantity and high-quality milk is that the health and welfare of the cows must be adequate. The association between herd size, milk production level, health, and welfare is complex and affected by many factors Chapinal et al. (2014a), including the managerial skills of the farmer, rate of herd expansion, facilities, training and experience of personnel, and the ratio of caretakers to animals. Barkema et al. (2015) provided evidence that larger farms in both the United States and China have a lower prevalence of lameness, corroborated by Chapinal et al. (2014b). However, according to Anastácio et al. (2014), herd-level and within-herd prevalence of infectious diseases in general

increase with increasing herd size (Wolf et al., 2014). This association may be confounded by recent herd expansion, including the purchase and mixing of animals from multiple sources, rather than being an effect of herd size itself. In many European countries, disease control programs have included the detection of DNA or antibodies in milk (Houe et al., 2006). Ensuring the well-being of Holstein cows is essential. Besides, the comprehensive animal health management protocols provide comfortable housing that allows for natural behaviours and promotes cow comfort, such as well-ventilated barns with adequate space, proper bedding, and clean water sources.

2.4 Manure and Waste Management

Holstein cows produce a large amount of manure, which can impact the environment if not managed properly. This waste is a significant source of N and P that, when land applied in excess of crop requirements, can cause contamination of surface water (Knowlton and Cobb, 2006). Excess N can also contaminate ground water through leaching. This poses a problem for human and animal health as consumed nitrate from drinking water is converted to nitrite in the digestive tract. One compound that affects air quality produced by dairy cattle is NH_3 . Ammonia is produced when N in urea from the animal's urine reacts with urease present in manure (Place and Mitloehner, 2010). A substantial GHG produced by dairy cattle waste is methane. The amount of CH_4 emitted by dairy waste is dependent on the amount of carbon, hydrogen, and oxygen present in the waste, making manure storage, diet, and bedding major contributors to total CH_4 production (Place and Mitloehner, 2010). Implementation of effective manure can help capture and utilise the nutrients in the manure, minimise water pollution, and reduce greenhouse gas emissions.

2.5 Energy Efficiency and Renewable Energy

It was revealed that production growth is the dominant contributor to the increase of GHG emissions (Kim and Kim, 2012), while changes in the energy mix, especially the contribution of renewable energy sources, reduce the GHG emissions (Marques et al., 2019). To maintain high milk quality, including low bacteria counts, milk cooling ensures a raw milk temperature of around 3-4 °C. Cooling systems are major energy users. Data are reported from 6.4 to 33.4 Wh/kg milk for CM (conventional milking systems) and 6.4 to 38.7 Wh/kg milk for AMS (automatic milking systems) (Upton et al., 2013). Warm water is required for technological needs, such as cleaning milking equipment, materials and buildings. The main systems for water heating are electric boilers or boilers heated by natural gas. Data range from 3.3 to 22.8 Wh/kg milk for electric boilers measured in Finnish dairy farms by Rajaniemi et al. (2017). Appropriate lighting can improve productivity and safety on a dairy farm. The average values for the contribution of lighting to the total milking process were 1.4 Wh/kg milk (Shine et al., 2018) and 32.1 Wh/kg milk for incandescent lamps (Houston et al., 2014). Besides the unit operations mentioned above, a number of other electricity uses are common in dairy farms. Data for miscellaneous energy users range from 4.1 to 38.8 Wh/kg milk. Despite ample evidence in the literature for the positive effects of solar panels on reducing fossil energy use, there is only partial support for the hypothesis, stating that "Dairy farms using solar panels are more energy efficient" (Houston et al., 2014). However, as it is fossil-fuel-generated energy use that is most pivotal for reducing greenhouse gas emissions, the generation of solar energy proves to be an important measure to make the sector more climate-proof, although attention should also be paid to the rebound effects of solar energy use, which are reflected in an increase in overall energy use (Qiu et al., 2019). At the same time, it is very important to reduce the environmental footprint of the dairy operation by implementing energy-efficient practices.

2.6 Water Conservation

Water is used in many aspects of dairy production within dairy barns. The total water footprint of the dairy industry makes up 19 % of the global water use of all animal production, second only to the beef industry at 33 % (Mekonnen and Hoekstra, 2012). However, it has been estimated that the overall water footprint of milk production is 1 m³/kg of milk (Mekonnen and Hoekstra, 2012). This represents a substantial amount of water that could be partially conserved throughout the production system on the farm (House et al., 2014). Robinson et al. (2016), based on their comprehensive study, came to the conclusion that free-stall dairy automated milking systems use more water on a daily basis than tie-stall and free-stall parlour operations. This leads to the reasoning that seasonality is a key factor in water use. Given that robotic facilities, which are becoming more common as milking systems, use a great amount of water, the industry should target efficient water-use strategies for these systems. Proper management of runoff and wastewater can also help prevent water contamination and protect local water sources.

2.7 Knowledge Sharing

To stay informed about sustainable practices, emerging technologies, and advancements in cow management, it is vital to engage with industry experts, researchers, and fellow farmers. Joining farmer networks or

organisations dedicated to sustainable agriculture allows for sharing experiences, learning from others, and collaboratively pursuing sustainability goals. Agricultural exhibitions, meetings, and events organised by herd-book organisations and farmers' associations serve as excellent venues for exchanging valuable information.

3. Conclusions

Responsible Holstein breeding necessitates a holistic approach, entailing the coordinated management of numerous factors impacting sustainability. The balanced management of the various elements highlighted in this study can significantly enhance the sustainability and resilience of Holstein-based dairy production. The application of genomic information and genomic selection schemes can also enable cows to produce more milk from less feed, minimising environmental impact. These strategies, combined with selective breeding for enhanced animal welfare, can lead to healthier cows with longer productive lives. Although the advantages of these approaches might not be immediately evident, their integration into a thorough management plan can greatly enhance the enduring viability and success of dairy farming.

References

- Anastácio S., Carolino N., Sidi-Boumedine K., Da Silva G.J., 2014, Q Fever Dairy Herd Status Determination Based on Serological and Molecular Analysis of Bulk Tank Milk. *Transboundary and Emerging Diseases*, 63, 293–300.
- Barkema H.W., Keyserlingk M.A.G. von, Kastelic J.P., Lam T.J.G.M., Luby C., Roy J.-P., 2015, Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J. of Dairy Science*, 98, 7426–7445.
- Beauchemin K.A., McAllister T.A., McGinn S.M., 2009, Dietary mitigation of enteric methane from cattle. *CAB Reviews*, 4, 1–18.
- Berry D.P., Bermingham M.L., Good M., More S.J., 2011, Genetics of animal health and disease in cattle. *Irish Veterinary Journal*, 31, 64–65.
- Brito L.F., Bedere N., Douhard F., Oliveira H.R., Arnal M., Peñagaricano F., Schinckel A.P., Baes C.F., Miglior F., 2021, Review: Genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world. *Animal*, 15, 100292, DOI: 10.1016/j.animal.2021.100292.
- Chapinal N., Liang Y., Weary D.M., Wang Y., Keyserlingk M.A.G. von, 2014a, Risk factors for lameness and hock injuries in Holstein herds in China. *Journal of Dairy Science*, 97, 4309–4316.
- Chapinal N., Weary D.M., Collins L., Keyserlingk M.A.G. von, 2014b, Lameness and hock injuries improve on farms participating in an assessment program. *Veterinary Journal*, 202, 646–648.
- Clay N., Garnett T., Lorimer J., 2020, Dairy intensification: Drivers, impacts and alternatives. *Ambio*, 49, 35–48.
- Cole J.B., VanRaden P.M., 2018, Symposium review: Possibilities in an age of genomics: The future of selection indices. *Journal of Dairy Science*, 101, 3686–3701.
- Doreau M., Rochette Y., Martin C., 2012, Effect of type of forage (maize silage vs grass silage) and protein source (soybean meal vs dehydrated lucerne) in dairy cow diet on methane emission and on nitrogen losses. In: Hassouna M., Guingand N. (Eds), *Proc. International Symposium on Emission of Gas and Dust from Livestock*, 10-13 June 2012, Saint-Malo, France, 4.
- Drake M.A., 2007, Invited review: Sensory analysis of dairy foods. *Journal of Dairy Science*, 90, 4925–4937.
- Gerber P.J., Henderson B., Makkar H.P. (Eds), 2013, *Mitigation of greenhouse gas emissions in livestock production: A review of technical options for non-CO₂ emissions*, FAO Animal Production and Health Papers. 177, Food and Agriculture Organization of the United Nations, Rome, Italy.
- Houe H., Lindberg A., Moennig V., 2006, Test strategies in bovine viral diarrhoea virus control and eradication campaigns in Europe. *Journal of Veterinary Diagnostic Investigation*, 18, 427–436.
- House H.K., Hawkins B.C., Barkes B.C., 2014, *Measuring and Characterizing On-Farm Milking Centre Washwater Volumes*. ASABE Paper, No 1908138, ASABE, 13-16 July 2014, Montreal, Quebec, Canada.
- Houston C., Gyamfi S., Whale J., 2014, Evaluation of energy efficiency and renewable energy generation opportunities for small scale dairy farms: A case study in Prince Edward residential solar panel adoption. *Journal of Environmental Economics and Management*, 96, 310–341.
- Hristov A.N., Lee C., Cassidy T., Heyler K., Tekippe J.A., Varga G.A., Corl, B., Brandt, R.C., 2013a, Effect of *Origanum vulgare* L. leaves on rumen fermentation, production, and milk fatty acid composition in lactating dairy cows, an inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production. *Journal of Dairy Science*, 96, 1189–1202.
- Hristov A.N., Oh J., Firkins J.L., Dijkstra J., Kebreab E., Waghorn G., Makkar, H.P.S., Adesogan, A.T., Yang W., Lee C., Gerber P.J., Henderson B., Tricarico J.M., 2013b, Special topics – Mitigation of methane and nitrous oxide emissions from animal operations: I. A review of enteric methane mitigation options. *Journal of Animal Science*, 91, 5045–5069.

- Hristov A.N., Oh J., Giallongo F., Frederick T.W., Harper M.T., Weeks H.L., Branco A.F., Moate P.J., Deighton M.H., Williams, S.R.O., Kindermann M., Duval S., 2015, An inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production. *Proceedings of the National Academy of Sciences*, 112, 10663–10668.
- Ivan M., Mir P., Mir Z., Entz T., He M.L., McAllister T.A., 2004, Effects of dietary sunflower seeds on rumen protozoa and growth of lambs. *British Journal of Nutrition*, 92, 303–310.
- Johnson D., Phetteplace H., Seidl A., 2002, Methane, nitrous oxide and carbon dioxide emissions from ruminant livestock production systems, In: Takahashi J., Young B.A., Soliva C.R., Kreuzer M. (Eds), *Greenhouse Gases and Animal Agriculture. Proceedings of the 1st International Conference on Greenhouse Gases and Animal Agriculture*, Obihiro, Japan, 7–11 November, 2001, Elsevier, Amsterdam, The Netherlands, 77–85.
- Johnson K.A., Johnson D.E., 1995, Methane emissions from cattle. *Journal of Animal Science*, 73, 2483–2492.
- Keyserlingk M.A.G. von, Martin N.P., Kebreab E., Knowlton K.F., Grant R.J., Stephenson M., Sniffen C.J., Harner III J.P., Wright A.D., Smith S.I., 2013, Invited review: Sustainability of the US dairy industry. *Journal of Dairy Science*, 96, 5405–5425.
- Kim K., Kim Y., 2012, International comparison of industrial CO₂ emission trends and the energy efficiency paradox utilizing production-based decomposition. *Energy Economics*, 34, 1724–1741.
- Knapp J.R., Laur G., Vadas P.A., Weiss W.P., Tricarico J.M., 2014, Invited review: Enteric methane in dairy cattle production: quantifying the opportunities and impact of reducing emissions. *Journal of Dairy Science*, 97, 3231–3261.
- Knowlton K., Cobb T.D., 2006, ADSA Foundation Scholar Award: Implementing waste solutions for dairy and livestock farms. *Journal of Dairy Science*, 89, 1372–1383.
- Lovett D., Shalloo L., Dillon P., O'Mara, F.P., 2006, A systems approach to quantify greenhouse gas fluxes from pastoral dairy production as affected by management regime. *Agricultural Systems*, 88, 156–179.
- Martin N.P., Russelle M.P., Powell J.M., Sniffen C.J., Smith S.I., Tricarico J.M., Grant R.J., 2017, Sustainable forage and grain crop production for the US dairy industry. *Dairy Science*, 100, 9479–9494.
- Marques A.C., Fuinhas J.A., Tomas C., 2019, Energy efficiency and sustainable growth in industrial sectors in EU countries: A non-linear ARDL approach. *Journal of Cleaner Production*, 239, 118045.
- McCaughy W., Wittenberg K., Corrigan D., 1999, Impact of pasture type on methane production by lactating beef cows. *Canadian Journal of Animal Science*, 79, 221–226.
- Mekonnen M.M., Hoekstra A.Y., 2012, A global assessment of the water footprint of farm animal products. *Ecosystems*, 15, 401–415.
- Miglior F., Fleming A., Malchiodi F., Brito L.F., Martin P., Baes C.F., 2017, A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *J. of Dairy Science*, 100, 10251–10271.
- Milani F., Nutter D. and Thoma G., 2011, Invited review: environmental impacts of dairy processing and products: a review. *J. Dairy Sci.* 94, 4243–4254.
- Moe P., Tyrrell H., 1979, Methane production in dairy cows. *Journal of Dairy Science*, 62, 1583–1586.
- Naranjo A., Johnson A., Rossow H., Kebreab E., 2020, Greenhouse gas, water, and land footprint per unit of production of the California dairy industry over 50 years. *Journal of Dairy Science*, 103, 3760–3773.
- National Research Council, 2001, *Nutrient Requirements of Dairy Cattle. Seventh Revised Edition*, The National Academies Press, Washington, DC, United States.
- Peterson C.B., Mitloehner F.M., 2021, Sustainability of the Dairy Industry: Emissions and Mitigation Opportunities. *Frontiers in Animal Science*, 2, No 760310, 1–14.
- Pinares-Patiño C., Baumont R., Martin C., 2003, Methane emissions by Charolais cows grazing a monospecific pasture of timothy at four stages of maturity. *Canadian Journal of Animal Science*, 83, 769–777.
- Qiu Y., Kahn M.E., Xing B., 2019, Quantifying the rebound effects of residential solar panel adoption. *Journal of Environmental Economics and Management*, 96, 310–341.
- Rajaniemi M., Jokiniemi T., Alakukku L., Ahokas J., 2017, Electric energy consumption on some Finnish dairy farms. *Agricultural and Food Science*, 26, 160–172.
- Robinson A.D., Gordon R.J., VanderZaag A.C., Rennie T. J., Osborne V.R., 2016, Usage and attitudes of water conservation on Ontario dairy farms. *The Professional Animal Scientist*, 32, 236–242.
- Shine P., Scully T., Upton J., Shalloo L., Murphy M.D., 2018, Electricity and direct water consumption on Irish pasture based dairy farms: A statistical analysis. *Applied Energy*, 210, 529–537.
- Upton J., Humpreys J., Groot Koerkamp P.W.G., French P., Dillon P., De Boer I.J.M., 2013, Energy demand on dairy farms in Ireland. *Journal of Dairy Science*, 96, 6489–6498.
- Wolf R., Barkema H.W., De Buck J., Slomp M., Flaig J., Hauptstein D., Pickel C., Orsel K., 2014, High herd-level prevalence of *Mycobacterium avium* subspecies *paratuberculosis* in Western Canadian dairy farms, based on environmental sampling. *Journal of Dairy Science*, 97, 6250–6259.