

ARTÍCULO CIENTÍFICO

CLASIFICACIÓN DE TOROS LECHEROS MEDIANTE ANÁLISIS DE FACTORES Y ANÁLISIS DE CONGLOMERADOS

Luis Alejandro Rodríguez-Campos^{1✉}

RESUMEN

El uso de semen de toros genéticamente seleccionados es ampliamente difundido en explotaciones lecheras, pero la gran cantidad de características por analizar puede hacer que la elección de un toro sea difícil. El presente trabajo buscó: 1) Determinar los factores subyacentes a las evaluaciones genéticas de toros lecheros. 2) Clasificar los toros en grupos con características similares. 3) Comparar los grupos generados con respecto a índices bioeconómicos y de tipo, número de hijas y precio de semen. Se usó un archivo de datos con 3021 toros Holstein evaluados en los Estados Unidos y se aplicó un análisis de factores con rotación oblimin. Con los factores extraídos se crearon conglomerados mediante el método de k-medias. Índices bioeconómicos (Mérito neto, mérito quesero, mérito fluido, mérito pastoreo, TPI), de tipo (PTA Tipo, compuesto de ubres, compuesto de patas), el número promedio de hijas en hatos de EEUU y el precio promedio de la dosis de semen fueron comparados mediante Kruskal-Wallis. Cinco factores que explicaron el 79% de la variabilidad: Estructura, Permanencia, Producción, Pezones y Aplomos. Mediante estos factores se encontraron siete grupos, que mostraron diferencias significativas ($p < 0,001$) en los diferentes índices bioeconómicos y de tipo, así como en el número de hijas pero el precio del semen mostró valores similares para la mayoría de los grupos. El uso de los cinco factores analizados podría facilitar la selección de toros de selección artificial, ayudando a perfilar más adecuadamente los animales.

Palabras clave: Ganado de leche, evaluación de sementales, mejoramiento animal, recursos genéticos animales

^{1✉} Universidad de Costa Rica, Estudiante de Maestría Académica en Estadística. San José, Costa Rica.
Autor para correspondencia: luisalejandro.rodriguez@ucr.ac.cr
Recibido: 15 abril 2018 Aceptado: 29 julio 2019



ABSTRACT

Dairy bull clasification using factor and cluster analyses. Using semen of genetically selected sires is a common practice in commercial dairy farms, however, there is a huge number of characteristics to consider, which complicates sire selection. This work aimed to 1) Determine underlying factors in genetic evaluation of dairy sires. 2) Classify sires on similar groups. 3) Compare these groups respecting to bioeconomic and type indexes, number of daughters and semen price. We used a database with information of 3021 sires evaluated at the USA, and applied a factor analysis, with oblimin rotation. Using the extracted factors, we created clusters, by k-means method. Bioeconomic (Net Merit, Cheese Merit, Fluid Merit, Grazing Merit, TPI), type (Type PTA, Udder Composite, Feet and Leg Composite) indexes, number of daughters in USA herds and mean semen price were compared between clusters using Kruskal-Wallis test. Five factors explained 79% of the data variability: Structure, Herd Permanence, Production, Teats and Feet and Leg Conformation. Using these factors we found seven clusters, with important differences in bioeconomic and type indexes, in addition to daughter numbers. However, semen price showed little inter-cluster variation. Using the five dimensions found could help to ease the process of artificial insemination sire selection.

Key words: Dairy cattle, sire evaluation, animal improvement, animal genetic resources

INTRODUCCION

La inseminación artificial es una técnica de manejo reproductivo ampliamente usada en las explotaciones lecheras y de doble propósito a nivel mundial, ya que permite reducir el número de toros que serían necesarios para mantener la reproducción del hato, además de acelerar la mejora genética, teniendo acceso incluso a material importado (TNAU, 2009).

Para la aplicación de esta biotecnología reproductiva, es ideal contar con semen de toros sanos, libres de enfermedades genéticamente transmisibles y que sean transmisores de superioridad genética para una o varias características productivas (Pallete, 2001). En este sentido, las autoridades competentes en cada país realizan las evaluaciones necesarias, basadas en el conocimiento de los sistemas productivos de cada país, y el establecimiento de una “meta de mejora” para seleccionar los animales más aptos para ser los progenitores de la siguiente generación (Oldenbroek y Van der Waaij, 2015). Posterior a esta fase, las compañías comerciales que poseen los toros se dedican a la comercialización del material genético, por medio de los llamados “catálogos de inseminación”, donde se provee una lista de animales, con los valores de las “Habilidades Transmisoras” para cada característica, más información auxiliar como nombre del toro, criador, entre otras.

Los catálogos de semen de compañías comerciales incluyen una gran cantidad de valores diferentes (hasta 70, incluyendo habilidades transmisoras y las confiabilidades de estas, índices combinados y descriptivos de la cantidad de hijas del toro) esto dificulta seleccionar los mejores animales sin ignorar ciertas características. Diferentes índices que combinan características de interés productivo han sido creados por las autoridades competentes y las asociaciones de criadores con el fin de dar un criterio único de selección, esto basado en análisis bioeconómicos (Oldenbroek y Van der Waaij, 2015). Sin embargo, la jerarquía de un toro en cada uno de estos índices es dependiente de los pesos que se le otorgan a cada variable, que pueden cambiar con el tiempo y limitar el avance en metas particulares del productor.

La hipótesis del presente trabajo es que hay un número pequeño de variables subyacentes que controlan todas las características medidas directamente en las evaluaciones genéticas. De poder cuantificar dichos factores, se podría facilitar la labor de selección de animales, haciendo más fáciles de consumir los diferentes catálogos de inseminación.

Basado en lo anterior, el presente trabajo tiene como objetivos los siguientes: 1) Determinar los factores subyacentes a los valores de habilidades transmisoras de los toros lecheros evaluados en los Estados Unidos. 2) Clasificar, con base en los factores encontrados, los toros en grupos con características similares. 3) Comparar los grupos generados con respecto a índices bioeconómicos y de tipo, número de hijas y precio de semen, para determinar si los factores encontrados permiten discriminar cuáles toros aportan un mayor rendimiento económico a las explotaciones.

MATERIALES Y MÉTODOS

Origen de los datos. Se utilizó la base de datos de evaluaciones genéticas de tipo, conformación y otros índices de mérito para toros en estado Activo, Limitado, Extranjero y Genómico en los Estados Unidos de América. Dicha base de datos es publicada por la Asociación Nacional de Criadores de Animales de EEUU (NAAB, por sus siglas en inglés) en los meses de abril, agosto y diciembre, a partir de los datos publicados por el Consejo de Reproducción del Ganado Lechero (CDCB) y el Programa de Mejoramiento Animal del Departamento de Agricultura de los Estados Unidos (USDA-AIPL). En este trabajo se utilizó un archivo descargado en octubre de 2016, pero la versión más reciente está disponible en la dirección <https://www.naab-css.org/database-files>.

Selección de las variables. El archivo original contiene 95 variables, la lista original está disponible en https://redmine.uscdcb.com/projects/cdcb-customer-service/wiki/Format_38. De esta lista original se preseleccionaron para el análisis todas las habilidades transmisoras predichas (PTAs por sus siglas en inglés). Sin embargo, las habilidades

transmisoras de movilidad (PTA Mobility) y tasa de concepción del toro (SCR) fueron excluidas por contar con una gran cantidad de valores perdidos (97% y 67%, respectivamente).

Se seleccionaron solamente los casos completos de animales pertenecientes de raza Holstein (negro y rojo) (n = 3021). La lista de variables con las que se trabajó y sus estadísticos descriptivos se muestra en el Cuadro 1. Las mismas fueron re-escaladas mediante normalización¹ en un rango de 0 a 10 para evitar el efecto de valores negativos y diferentes escalas de medición. Además, las variables PCS y VLPT fueron invertidas², para favorecer cargas positivas.

Se realizó el análisis de los indicadores de adecuación muestral individuales (prueba de Kaiser-Meyer-Olkin), descartando a todas aquellas variables que presentaran valores menores a 0,5. Los resultados de este cálculo se muestran también en el Cuadro 1. Se debe mencionar que el valor global de adecuación muestral es de 0,75, que se considera apropiado (Kaiser y Rice, 1974).

¹ Se denomina normalización a la transformación $Y_i = (X_i - \text{Min}(X))/(\text{Max}(X)-\text{Min}(X))$

² La inversión se realizó mediante la siguiente transformación: $Y_i = (X_i - \text{Max}(X))/(\text{Min}(X)-\text{Max}(X))$

Cuadro 1. Estadísticos descriptivos de las habilidades transmisoras predichas y habilidades transmisoras estandarizadas disponibles en la evaluación genética de toros de EEUU, utilizadas en el presente estudio (n = 3021).

Variable	Significado	Min	Media	Max	D.E	MSAi
ANG	Angulo de la grupa	-4,2	0,0	4,5	0,9	0,46
ANI	Ancho de los isquiones	-3,5	0,8	4,1	0,9	0,93
ANUT	Ancho de la ubre trasera	-2,9	2,0	4,6	1,1	0,84
APEZ	Angulo de la pezuña	-3,2	1,2	4,8	0,9	0,85
AUT	Altura de la ubre trasera	-3,0	2,2	5,0	1,2	0,84
CGR	Concentración (%) de grasa en la leche	-0,3	0,1	0,4	0,1	0,43
CPRO	Concentración (%) de proteína en la leche	-0,2	0,0	0,2	0,0	0,37
EST	Estatura	-5,2	1,1	4,8	1,1	0,78
FL	Forma Lechera	-4,9	0,9	4,2	1,0	0,68
FLSM	Fortaleza del ligamento suspensorio medial de la ubre	-3,6	1,0	3,7	0,9	0,83
FOR	Fortaleza	-3,3	0,6	3,8	0,8	0,67
FPH	Facilidad de parto en las hijas	1,0	5,8	13,3	1,5	0,84
FPT	Facilidad de parto del toro	3,7	7,2	13,2	1,2	0,73
IDU	Inserción delantera de la ubre	-3,1	1,7	4,5	1,1	0,81
LIV	Livability	-6,2	0,9	7,3	1,9	0,80
LPEZ	Largo de los pezones	-4,4	-0,2	3,8	0,9	0,75
NVH	Tasa de nacidos vivos en las hijas	2,6	6,2	13,7	1,5	0,82
NVT	Tasa de nacidos vivos del toro	4,6	7,5	10,7	0,7	0,80
PC	Profundidad Corporal	-3,9	0,5	3,8	0,9	0,77
PCS	Puntaje de células somáticas	2,4	2,8	3,5	0,1	0,77
PGR	Producción de grasa en leche (lb)	-71,0	48,4	125,0	26,9	0,64
PLE	Producción de leche (lb)	-2815	923	2870	670	0,56
PPRO	Produccion de proteínas lácteas (lb)	-71,0	35,1	87,0	18,0	0,65
PRU	Profundidad de la ubre	-3,2	1,3	4,3	1,0	0,70
TCN	Tasa de concepción en novillas	-5,9	1,5	6,0	1,4	0,89
TCV	Tasa de concepción en vacas	-8,5	1,8	7,6	2,2	0,79
TPH	Tasa de preñez en las hijas	-5,5	1,0	6,4	1,7	0,78
UPA	Ubicación de los pezones anteriores	-2,5	0,8	3,7	0,8	0,74
UPP	Ubicación de los pezones posteriores	-3,6	0,8	4,0	0,9	0,67
VLPT	Vista lateral de las patas traseras	-3,9	-0,2	3,5	0,9	0,61
VP	Vida Productiva	-5,4	4,0	10,9	2,4	0,82
VTPT	Vista trasera de las patas traseras	-3,9	1,2	4,6	0,9	0,89

Min = Mínimo, Max = Máximo, 75, D.E. = Desviación estándar, MSAi = Índice de adecuación muestral

Análisis de factores. Para determinar el número de factores subyacentes por extraer se comparó el resultado de varias metodologías (autovalores mayores a 1, análisis paralelo, estructura muy simple, el criterio MAP de Velicer). Dichas metodologías coincidieron en sugerir la extracción de cinco factores.

Los factores subyacentes del conjunto de datos resultante se extrajeron por el método de componentes principales, aplicando luego la rotación oblimin. Para ello se usó la función `principal()` del paquete `psych` versión 1.6.4 (Revelle, 2016), complemento del software estadístico R, versión 4.3 (R Core Team, 2017).

Se hicieron descartes sucesivos eliminando aquellas variables al descartar aquellas variables con cargas significativas en más de un factor (complejidades mayores o iguales a 2). En este paso las variables descartadas fueron FPT, FPH, NVT, NVH, IUD, AUT, ANUD, PU y LP. En el Cuadro 2 se presentan los coeficientes de correlación de Pearson entre las variables incluidas en el modelo final.

Análisis de conglomerados. Con los puntajes factoriales obtenidos se procedió a realizar un análisis de conglomerados, con el fin de formar grupos de individuos con similares características, a partir del método de k-medias. Para ello se usó la función `kmeans()` de R 4.3 (R Core Team, 2017). Para determinar el número apropiado de grupos para formar, se siguió una modificación a la metodología del codo (Kassambara, 2017): se realizaron 100 repeticiones del algoritmo k-medias, con diferentes puntos de arranque, extrayendo las sumas de cuadrados entre grupos, para valores de k de 1 a 25. Se graficó el promedio de la varianza entre grupos con cada k (Figura 1A). Posteriormente se calculó el cambio porcentual en la suma de cuadrados entre grupos en cada una de las repeticiones y se promedió dicho valor en las 100 repeticiones. Se graficó el cambio porcentual (Figura 1B) y se usó como el número de clústeres adecuado el primer valor que arrojara un cambio porcentual en la varianza entre grupos menor a 5% ($k = 7$).

Cuadro 2. Coeficientes de correlación de Pearson (valores encima de la diagonal) y significancia de dichas correlaciones (valores bajo la diagonal) para las variables analizadas.

	PLE	PGR	PPRO	PCS	VP	TPH	TCN	TCV	LIV	EST	FOR	PC	ANI	VLPT	VTPT	APEZ	FLSM	UPA	UPP
PLE	-	0,58	0,87	-0,01	0,35	-0,04	0,17	0,18	0,10	0,06	0,04	0,08	0,12	0,02	0,14	0,06	0,05	0,04	0,04
PGR	<0,01	-	0,74	-0,15	0,47	0,07	0,30	0,28	0,13	0,08	0,10	0,14	0,16	-0,09	0,28	0,19	0,13	0,10	0,09
PPRO	<0,01	<0,01	-	-0,05	0,43	0,06	0,20	0,27	0,14	0,06	0,06	0,07	0,11	0,01	0,17	0,09	0,04	0,05	0,04
PCS	0,71	<0,01	<0,01	-	-0,49	-0,35	-0,26	-0,35	-0,25	-0,09	-0,12	0,01	-0,05	0,08	-0,13	-0,21	-0,09	-0,05	0,01
VP	<0,01	<0,01	<0,01	<0,01	-	0,65	0,55	0,77	0,69	-0,23	-0,21	-0,33	-0,12	-0,08	0,16	0,10	0,01	0,10	0,01
TPH	0,04	<0,01	<0,01	<0,01	<0,01	-	0,57	0,90	0,48	-0,17	-0,17	-0,34	-0,17	-0,07	-0,05	0,01	-0,09	0,02	-0,06
TCN	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,70	0,28	0,07	-0,03	-0,05	0,13	0,00	0,17	0,14	0,16	0,14	0,10
TCV	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,54	-0,19	-0,20	-0,33	-0,12	-0,04	-0,03	-0,01	-0,05	0,05	-0,02
LIV	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	-0,35	-0,31	-0,42	-0,28	0,07	-0,17	-0,22	-0,11	-0,04	-0,05
EST	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,76	0,83	0,74	0,04	0,45	0,58	0,55	0,25	0,24
FOR	0,03	<0,01	<0,01	<0,01	<0,01	<0,01	0,11	<0,01	<0,01	<0,01	-	0,90	0,69	-0,14	0,49	0,59	0,33	0,14	0,11
PC	<0,01	<0,01	<0,01	0,70	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,74	0,03	0,46	0,51	0,46	0,23	0,24
ANI	<0,01	<0,01	<0,01	0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,13	0,40	0,42	0,52	0,31	0,32
VLPT	0,26	<0,01	0,72	<0,01	<0,01	<0,01	0,93	0,01	<0,01	0,02	<0,01	0,09	<0,01	-	-0,43	-0,53	0,18	0,12	0,18
VTPT	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	0,14	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,79	0,33	0,24	0,15
APEZ	<0,01	<0,01	<0,01	<0,01	<0,01	0,58	<0,01	0,53	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,31	0,19	0,09
FLSM	0,01	<0,01	0,02	<0,01	0,46	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,64	0,78
UPA	0,04	<0,01	0,01	0,01	<0,01	0,20	<0,01	0,01	0,02	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	0,87
UPP	0,02	<0,01	0,02	0,44	0,46	<0,01	<0,01	0,41	0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-

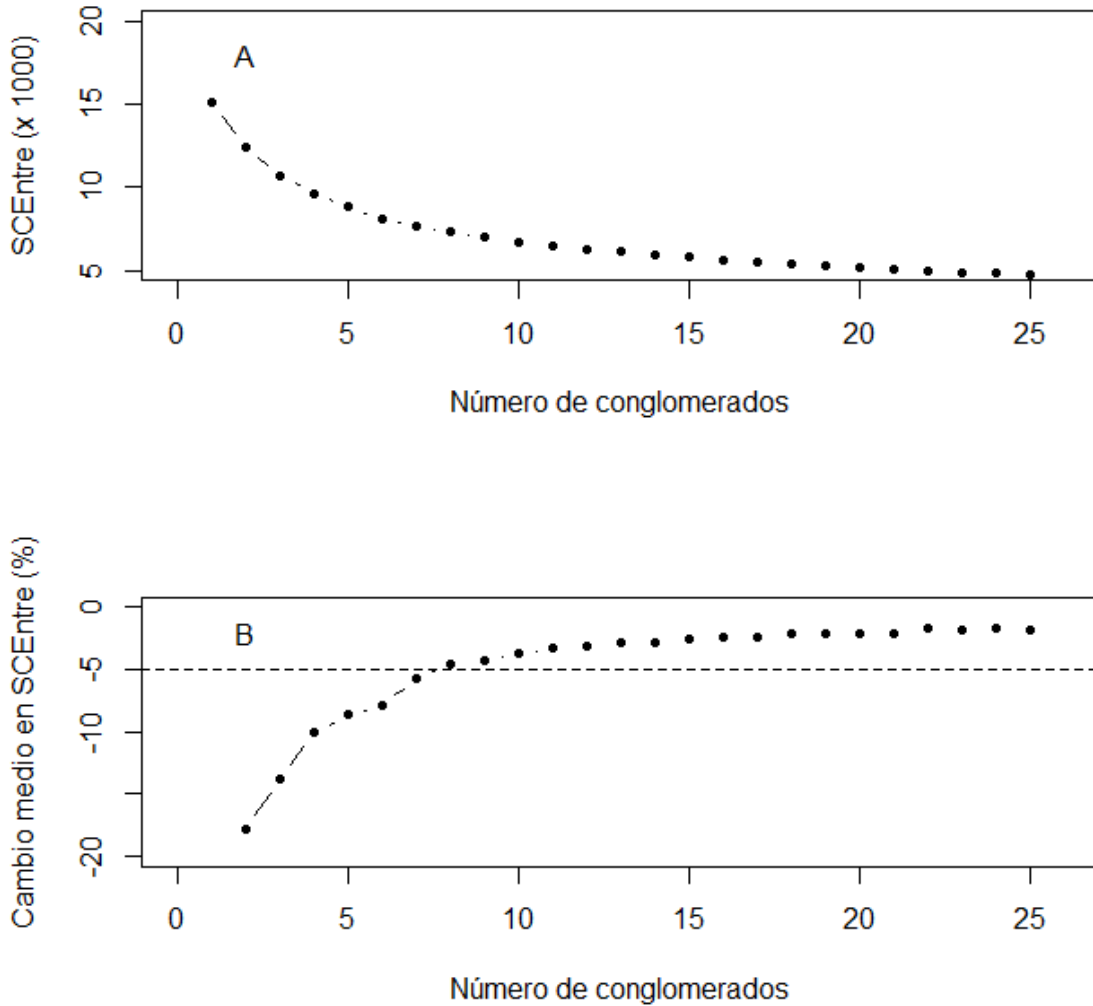


Figura 1. Criterio gráfico para la selección del número de conglomerados, a partir de 100 repeticiones. Panel A) Suma de cuadrados entre grupos (SCEntre) dependiendo del número de conglomerados. Panel B) Cambio porcentual promedio en la SCEntre por cada conglomerado adicional.

Posterior a ello se obtuvo los centroides de los conglomerados y se validó el resultado con los promedios y desviaciones estándares de diferentes índices de productividad

(Mérito Fluido, Mérito Quesero, Mérito Pastoreo, Mérito Neto, PTA de Tipo, Índice de Tipo y Rendimiento (TPI), Índice Compuesto de Ubres e Índice Compuesto de Patas), el número de Número de Hijas en EEUU y el precio por dosis de semen. Estos valores se compararon usando la prueba de Kruskal-Wallis (dado que se evidenciaron violaciones al supuesto de homocedasticidad), y las comparaciones de a pares se realizaron mediante el ajuste de Holm. Para ello se usó la función `ea1()` del paquete `Easynova` (Arnhold, 2013), de R.

RESULTADOS Y DISCUSIÓN

Factores subyacentes: Los 5 factores extraídos, sus cargas factoriales y la proporción de varianza explicada por cada uno se muestran en el Cuadro 3. Otros autores han encontrado soluciones factoriales con 4 (Chu et al., 2002), 7 (Corrales et al., 2011) y 2 factores (Kern et al., 2014) al analizar características de tipo en ganado Holstein en China, Colombia y Brasil, respectivamente. Todas estas son soluciones posibles, dependiendo de las variables que se incluyan. En pasos previos del análisis, los diferentes criterios de selección del número de factores arrojaron recomendaciones de extraer 2, 4, 5, 6 y 7 factores, dependiendo del criterio. Sin embargo, al analizar las cargas, se encontró que había variables con cargas fuertes en dos y hasta tres factores, así como variables que cargan solas en un factor (datos no mostrados). Este tipo de situaciones no deseables, se ven reflejadas en los trabajos de los tres grupos mencionados. Por otro lado, muchas de las variables contempladas en dichos análisis no se consideraron en el presente trabajo, así como en los trabajos de los autores mencionados no se consideran variables reproductivas y de producción en el análisis factorial.

El primer factor explica un 21% de la varianza total y está compuesto por las variables EST, PC, FOR y ANI. Dichas variables son indicadores morfológicos de la estructura ósea del animal, y cuán adaptado está para las exigencias del sistema de producción de leche. Similares resultados obtuvieron Corrales et al. (2011) al analizar la asociación

entre variables de tipo y producción en ganado Holstein de la región de Antioquia, Colombia. El primer factor extraído, denominado Estructura, incluyó el tamaño, la estatura, profundidad del cuerpo, ancho de isquiones y angularidad (sinónimo de forma lechera). Por su parte, Chu et al. (2002) encontró que variables como EST, PC, FOR y ANI cargaban en el mismo factor con las variables de sistema mamario como FLSM, PRU y UPP. En el presente conjunto de datos, previo a la eliminación de variables de alta complejidad, sucedió la misma situación. Aún así, se observa como la variable FLSM carga con 0,30 en el factor 1, coincidiendo con Chu et al. (2002). Esta relación entre la estructura ósea y el sistema mamario se ve reforzada por la correlación de 0,3 entre el factor 1 y el factor 4, lo cual no es posible de observar en los trabajos mencionados, ya que usan una rotación ortogonal (varimax).

El segundo factor está compuesto por las variables relacionadas con características reproductivas (TPH, TCV y TCN), salud (PCS) y vida productiva (VP, LIV). Se ha denominado permanencia, pues valores más altos en este factor aseguran vacas que tendrán las características suficientes para permanecer muchos años en el hato, proveyendo un buen retorno a la inversión del período de levante. Otros autores también han señalado la importancia de factores reproductivos y sanitarios, específicamente del conteo de células somáticas, en la toma de decisiones de descarte de los productores. Por ejemplo, Ansari-Lari et al. (2012) al analizar una población de ganado Holstein en Irán, detectó que el 33% de los descartes de vacas se debió a infertilidad, mientras que el 6% se debió a mastitis (enfermedad caracterizada por inflamación e infección de la ubre, de la cual el conteo de células somáticas es un indicador). De manera similar, Hadley et al. (2006), al analizar más de 7 millones de registros de descarte de vacas en los Estados Unidos, detectaron que el 19% de las ocasiones el descarte se debió a problemas reproductivos y el 12% a mastitis.

El tercer factor es denominado producción e incluye los volúmenes de producción de leche, grasa y proteína en lactancias corregidas a 305 días. Esto era de esperarse, pues dichas características están ampliamente correlacionadas, ya que animales con mayor producción láctea, generalmente darán más volumen de grasa y proteína que animales con baja producción pero alta concentración de sólidos. Curiosamente, las variables de concentración fueron excluidas del análisis, indicando que pertenecen a otro espacio vectorial.

Cuadro 3. Cargas factoriales, autovalores, porcentaje de varianza explicada y correlaciones entre factores para el modelo factorial analizado.

Cargas Factoriales:	Factor 1: Estructura	Factor 2: Permanencia	Factor 3: Producción	Factor 4: Pezones	Factor 5: Aplomos	Comunali- dad	Unicidad	Complejidad
EST	0,90	0,01	-0,03	0,06	0,01	0,85	0,150	1,0
PC	0,90	-0,17	0,06	0,01	0,02	0,91	0,091	1,1
FOR	0,90	-0,01	-0,01	-0,12	0,15	0,84	0,161	1,1
ANI	0,85	0,04	0,07	0,13	-0,12	0,77	0,228	1,1
TPH	-0,03	0,92	-0,19	-0,06	-0,01	0,82	0,182	1,1
TCV	-0,05	0,92	0,06	-0,02	-0,04	0,88	0,119	1,0
VP	-0,18	0,76	0,30	0,07	0,12	0,88	0,125	1,5
TCN	0,19	0,75	0,06	0,07	-0,03	0,60	0,404	1,2
PCS-	0,23	0,61	-0,12	-0,07	0,11	0,38	0,616	1,5
LIV	-0,28	0,61	0,07	0,00	-0,14	0,55	0,451	1,5
PPRO	0,01	0,02	0,96	-0,04	-0,02	0,92	0,080	1,0
PLE	0,00	-0,07	0,93	-0,03	-0,05	0,84	0,165	1,0
PGR	0,04	0,11	0,80	0,04	0,12	0,73	0,268	1,1
UPP	-0,08	-0,06	-0,01	0,99	-0,04	0,94	0,060	1,0
UPA	-0,08	0,02	-0,03	0,95	0,07	0,86	0,141	1,0
FLSM	0,30	0,03	-0,02	0,77	-0,03	0,81	0,188	1,3
VLPT-	-0,26	-0,05	-0,03	-0,11	0,92	0,82	0,176	1,2
APEZ	0,40	0,08	-0,01	0,07	0,74	0,88	0,115	1,6
VTPT	0,27	0,01	0,12	0,17	0,71	0,78	0,221	1,5
Autovalores	3,94	3,71	2,65	2,65	2,11	-	-	-
Varianza Explicada	21%	20%	14%	14%	11%	-	-	-
Correlación con:								
Permanencia	-0,17	-	-	-	-	-	-	-
Producción	0,08	0,22	-	-	-	-	-	-
Sistema Mamario	0,31	0,04	0,09	-	-	-	-	-
Aplomos	0,24	0,07	0,08	0,00	-	-	-	-

Variables marcadas con un guion (-) denotan variables con escala reversada. Valores en negrita son significativos.

El cuarto factor encontrado tiene que ver con la posición de los pezones, ya que se incluyen las variables UPA, UPP y FLSM. Las dos primeras variables miden el ángulo de desvío que presentan los pezones con respecto al vientre de la vaca. Por su parte, la

FSLM tiene una marcada influencia en la resistencia de la ubre ante el estrés mecánico producido por la acumulación de la secreción láctea en el período inter-ordeño, pero también se asocia con la posición de los pezones, pues un ligamento suspensorio de la ubre muy fuerte puede jalar los pezones hacia adentro, haciendo un ángulo muy cerrado. Dicha asociación se muestra en los trabajos de Chu et al. (2002) y Corrales et al. (2011).

El quinto factor encontrado corresponde a los aplomos del animal, evaluado como el ángulo de los corvejones, tanto en su vista lateral como trasera, y el ángulo de la pezuña. Es de notar que la variable VLPT debió ser reversada para mantener la consistencia del factor. Esta variable evalúa el ángulo del corvejón del animal, visto desde lateral. Aunque en esta variable, ambos extremos biológicos son indeseables, Weigel et al. (2003) indican que vacas con patas traseras muy rectas son 4% más propensas a descarte voluntario que aquellas con ángulos del corvejón intermedios, pero al disminuir el ángulo del corvejón esta propensión relativa aumenta hasta un 30%. Del mismo modo, puntajes bajos en ángulo de pezuña y vista lateral de patas traseras aumentan de un 10 a un 20% el riesgo de descarte (Weigel et al., 2003; Holstein Association, 2016). En conclusión, mayores puntajes en este factor se asocian a animales con mejor conformación de patas, lo cual es importante considerando que el 4% de los descartes de animales pueden deberse a cojeras (Hadley et al., 2006).

A pesar de lo anterior, no se determinó correlación entre este factor y la permanencia en el hato ($r = 0,07$), pero sí con la estructura ósea ($r = 0,24$), por cuanto animales hay correlaciones mayores a 0,4 entre EST, FOR, PC y ANI con VTPT y APEZ (Cuadro 2). Chu et al. (2002) encontraron un factor en el que FOR y APEZ cargaban juntos, mientras que VLPT y EST cargaban juntas en otro. Esto refuerza la interrelación entre estas variables.

Análisis de conglomerados. A partir del análisis de la Figura 1, se determinó la conveniencia de extraer siete grupos a partir de los puntajes factoriales que se obtuvo. La suma de cuadrados residual (entre grupos) corresponde a un 49,6% de la varianza total, con lo cual se concluye que la agrupación realizada explica poco más de la mitad de la varianza.

El Cuadro 4 muestra los centroides de cada uno de los conglomerados obtenidos. El grupo 1, destaca entre los demás por sus altos valores en el factor de permanencia, si bien es relativamente mediocre en los otros componentes. Lo anterior hace que sea denominado *Durables*, porque son toros que dan lugar a vacas con buenas características reproductivas, de salud y vida productiva. Este grupo, a pesar de que no es tan productivo, destaca por sus valores en Mérito de Pastoreo (Cuadro 5), lo cual es de esperarse, ya que este índice le da 43% de su peso a las variables de salud y reproducción, en comparación con los Méritos neto, Fluido o Quesero, para los cuales este componente tiene un peso de 33 a 37% (Van Raden, 2014). El número de hijas y el precio del semen también son superiores a otros grupos (Cuadro 5).

El grupo 2 corresponde a animales con importantes puntajes en el factor de producción, por ello se le denominó *Alto Volumen* (Cuadro 4). Estos animales presentan valores importantes en PLE, PPRO y PGR, lo cual da origen a valores altos en los índices de Mérito Fluido, Mérito Quesero, Mérito en Pastoreo y Mérito Neto, que dan una importancia a las características productivas de 37 a 42%. Esto también se ve reflejado en altos precios de semen e hijas en EEUU (Cuadro 5).

El grupo 3 está compuesto por animales denominados correctos: presentan valores intermedios para la mayoría de los factores, aunque presentan el mayor promedio del factor pezones, indicando que son animales cuya progenie suele ser de pezones encontrados (Cuadro 4). Estos animales destacan por su buena conformación, por lo cual, obtienen los valores más altos en los índices relacionados con características de tipo (PTA Tipo y compuesto de ubres). Dentro de los índices generales son los animales con puntajes más altos en el TPI (índice de tipo y rendimiento). Este índice, creado por la Asociación de Criadores de Ganado Holstein de los EEUU, le da una importancia de 28% a las características de tipo y conformación (Holstein USA, 2017), superior a otros índices como mérito neto, mérito quesero, mérito fluido y mérito en pastoreo, donde las variables de conformación no superan el 21% (Van Raden, 2017). Las características antes mencionadas hacen que sea el grupo con mayor precio de semen y más hijas por hato (probablemente hijas de estos toros son las grandes campeonas en competencias de ganado Holstein), aunque no significativamente mayor a los grupos anteriormente mencionados.

Cuadro 3. Centroides y tamaño de los siete conglomerados obtenidos mediante el procedimiento k-medias.

Factor	Conglomerado (tamaño)						
	Durables (528)	Alto Volumen (545)	Correctos (511)	Indeseables (286)	Exhibición (288)	Rectos (511)	Cerrados (352)
Estructura	-0,62	-0,07	0,80	-1,08	1,23	0,17	-0,51
Permanencia	1,14	0,10	0,22	-0,93	-1,31	-0,01	-0,35
Producción	-0,22	0,95	0,48	-1,29	-1,08	0,29	-0,30
Pezones	-0,02	-0,63	0,87	-1,46	0,54	-0,13	0,68
Aplomos	0,03	-0,35	0,12	-0,44	0,26	1,25	-1,35

Los datos están estandarizados.

Cuadro 4. Índices de mérito, número de hijas y precio de semen (Promedio \pm Error Estándar) según los conglomerados encontrados para diferentes índices de mérito usados en la evaluación genética de los Estados Unidos y número de hijas

Variable	Durables	Alto Volumen	Correctos	Indeseables	Exhibición	Aplomados	Cerrados
Mérito Fluido	559 \pm 0.2 c	627 \pm 0.2 a	585 \pm 0.2 b	126 \pm 0.7 e	171 \pm 0.6 e	556 \pm 0.3 c	368 \pm 0.5 d
Mérito Quesero	636 \pm 0.3 bc	680 \pm 0.2 a	654 \pm 0.2 b	176 \pm 0.7 e	221 \pm 0.7 e	617 \pm 0.3 c	410 \pm 0.5 d
Mérito Pastoreo	621 \pm 0.2 a	628 \pm 0.2 a	611 \pm 0.2 a	149 \pm 0.7 d	175 \pm 0.6 d	571 \pm 0.3 b	375 \pm 0.5 c
Mérito Neto	613 \pm 0.3 bc	665 \pm 0.2 a	634 \pm 0.2 b	161 \pm 0.7 e	206 \pm 0.6 e	599 \pm 0.3 c	398 \pm 0.5 d
PTA de Tipo	1.26 \pm 0.001 e	1.46 \pm 0.001 d	2.28 \pm 0.001 a	0.10 \pm 0.003 g	2.19 \pm 0.003 b	1.88 \pm 0.001 c	1.12 \pm 0.002 f
TPI	2394 \pm 0.3 d	2459 \pm 0.3 b	2498 \pm 0.3 a	1792 \pm 0.8 g	2005 \pm 0.8 f	2424 \pm 0.3 c	2121 \pm 0.6 e
Compuesto de Ubres	1.39 \pm 0.001 a	1.16 \pm 0.001 ab	2.00 \pm 0.001 b	0.13 \pm 0.003 c	1.85 \pm 0.003 d	1.74 \pm 0.001 d	1.13 \pm 0.002 e
Compuesto de Patas	1.03 \pm 0.001 c	0.92 \pm 0.001 c	1.56 \pm 0.001 b	0.26 \pm 0.003 d	1.59 \pm 0.003 b	1.88 \pm 0.001 a	0.22 \pm 0.002 d
Hijas en EEUU	96.7 \pm 0.03 a	96.8 \pm 0.03 a	98.8 \pm 0.02 a	53.2 \pm 0.17 c	79.4 \pm 0.12 b	96.1 \pm 0.03 a	81.8 \pm 0.10 b
Precio del Semen	28.6 \pm 0.02 a	31.8 \pm 0.02 ab	32.4 \pm 0.03 ab	20.7 \pm 0.03 ab	25.1 \pm 0.04 ab	30.7 \pm 0.02 ab	23.1 \pm 0.02 b

^{a,b} Grupos con letras diferentes difieren en sus rangos por la prueba de Kruskal-Wallis, con ajuste de Holm ($p < 0,001$)

El grupo 4 está conformado por animales denominados indeseables, por mostrar valores bajos en la mayoría de los factores (Cuadro 4). Esta característica hace que los valores en los índices combinados sean bajos y, que, en consecuencia, presenten una de las menores cantidades de hijas de todos los grupos (Cuadro 5). No obstante, el precio de su semen no es significativamente diferente a los grupos 1, 2, 3, 4, 5 y 6.

El grupo 5, está conformado por animales que presentan valores extremos en las características de estructura, como EST, PC, FOR y ANI, y valores intermedios en otras características de tipo como el factor Pezones y el factor Aplomos, sin embargo, sus características reproductivas, productivas y de salud son deficientes (Cuadro 4). Se denomina a este grupo "Exhibición", ya que probablemente sean animales correctos que ganen concursos de tipo, pero no son rentables a nivel productivo. Esto se demuestra en sus valores mediocres para todos los índices bio-económicos evaluados (mérito neto, mérito quesero, mérito fluido, mérito pastoreo, TPI), pero valores buenos en los índices de tipo y conformación (Compuesto de Ubres, Compuesto de Patas, PTA de Tipo). La cantidad promedio de hijas indica que no es el favorito del productor estadounidense, pero tampoco es despreciado.

Los grupos 6 y 7 son extremos biológicos para las variables relacionadas con el factor aplomos y pezones (Cuadro 4). El primer grupo (Rectos) tiende a presentar patas rectas, con poco ángulo en la pezuña y corvejón. Esto provoca que sean animales que pueden tener problemas importantes de movilidad, debido al impacto sobre las articulaciones (Weigel et al., 2013). Sin embargo, sus valores de índice compuesto de patas son superiores a otros grupos. El grupo 7 (cerrados) presenta los valores de ángulo de pezones más cerrados, lo cual es inadecuado para el ordeño automatizado. Este último grupo presenta el valor promedio de semen más bajo de todos (Cuadro 5).

CONSIDERACIONES FINALES

El análisis presentado anteriormente permite comprobar la alta variabilidad entre los animales incluidos en la evaluación genética de los Estados Unidos. El uso del análisis de factores permite simplificar la elección de los animales, reduciendo la cantidad de variables por considerar. La agrupación en conglomerados permite observar cuáles toros presentan los valores más aptos para los diferentes intereses productivos, permitiendo obtener un perfil más holístico de los mismos.

El uso conjunto de ambas técnicas puede facilitar la importación de semen en países como Costa Rica, que dependen en gran medida del material genético evaluado en otros países, así como disminuir la complejidad en la elección de los padrotes a nivel de finca.

LITERATURA CITADA

- Ansari-Lari, M., M. Mohebbi-Fani, A. Rowshan-Ghasrodashti. 2012. Causes of culling in dairy cows and its relation to age at culling and interval from calving in Shiraz, Southern Iran. *Veterinary Research Forum* 3(4):233–237.
- Arnhold, E. 2013. Package in the R environment for analysis of variance and complementary analyses. *Brazilian Journal of Veterinary Research and Animal Science* 50(6):488-492.
- Chu, M.X., S.K. Shi. 2002. Phenotypic factor analysis for linear type traits in Beijing Holstein cows. *Asian-Australian Journal of Animal Science* 15(11): 1527-1530.
- Corrales, J. A., M. Cerón-Muñoz, J. A. Cañas, C. R. Herrera, S.C. Calvo. 2011. Relationship between type traits and milk production in Holstein cows from Antioquia, Colombia. *Revista MVZ Córdoba* 16:2507-2513.

- Hadley, G.L., C. A. Wolf, S. B. Harsh. 2006. Dairy cattle culling patterns, explanations, and implications. *Journal of Dairy Science* 89:2286–2296.
- Holstein USA (Asociación de criadores de Ganado Holstein de EEUU). 2017. Updates to the Total Performance Index® (TPI®) and Type Composites: Keeping up with the changing times. Consultado: 8 abril 2018. Disponible en: http://www.holsteinusa.com/pdf/Upcoming_Changes_aug17.pdf
- Kaiser, H. F., J. Rice. 1974. Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1):111–117.
- Kern, E. L., J.A. Cobuci, C.N. Costa, C.M.M. Pimentel. 2014. Factor analysis of linear type traits and their relation with longevity in Brazilian Holstein cattle. *Asian-Australasian Journal of Animal Sciences*, 27(6), 784–790.
- Oldenbroek, K., L. Van der Waaij. 2015. Textbook animal breeding and genetics for Bsc students. Centre for Genetic Resources the Netherlands and Animal Breeding and Genomics Centre. 311 p.
- Pallete, A. 2001. Evaluación y selección de toros lecheros. *Revista de Investigación Veterinaria del Perú* 12(2): 150-160.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. En: <https://www.R-project.org/>.
- Revelle, W. 2017. Package Psych: Procedures for Psychological, Psychometric, and Personality Research. R package documentation. En: <https://cran.r-project.org/web/packages/psych/psych.pdf>
- Van Raden, P.M. 2017. Net merit as a measure of lifetime profit: 2017 revision. Animal Improvement Program, Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD. En: <https://www.aipl.arsusda.gov/reference/nmcalc-2017.htm>

Weigel, K.A., R.W. Palmer, D.Z. Caraviello. 2003. Investigation of Factors Affecting Voluntary and Involuntary Culling in Expanding Dairy Herds in Wisconsin using Survival Analysis. *Journal of Dairy Science* 86:1482 – 1486.