



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Lactation curve models for estimating gene effects over a timeline

**Citation for published version:**

Strucken, EM, de Koning, DJ, Rahmatalla, SA & Brockmann, GA 2011, 'Lactation curve models for estimating gene effects over a timeline' Journal of Dairy Science, vol. 94, no. 1, pp. 442-449. DOI: 10.3168/jds.2009-2932

**Digital Object Identifier (DOI):**

[10.3168/jds.2009-2932](https://doi.org/10.3168/jds.2009-2932)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Journal of Dairy Science

**Publisher Rights Statement:**

© American Dairy Science Association®, 2011.

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





## Lactation curve models for estimating gene effects over a timeline

E. M. Strucken,\*† D. J. de Koning,† S. A. Rahmatalla,\* and G. A. Brockmann\*<sup>1</sup>

\*Breeding Biology and Molecular Genetics, Department for Crop and Animal Sciences, Faculty of Agriculture and Horticulture, Humboldt-Universität zu Berlin, Invalidenstrasse 42, D-10115 Berlin, Germany

†Roslin Institute and R(D)SVS, University of Edinburgh, EH25 9PS Roslin, United Kingdom

### ABSTRACT

The effects of genes are commonly estimated using random regression models based on test-day data and only give a general gene effect. Alternatively, lactation curve models can be used to estimate biological and environmental effects, or to predict missing test-day data and perform breeding value estimation. This study combines lactation curve models and estimation of gene effects to represent gene effects in different stages of lactation. The lactation curve models used were based on the Wood, Wilmink, and Ali and Schaeffer models. A random regression test-day model was used to compare estimated gene effects with the results of commonly used models. The well-characterized *DGAT1* gene with known effects on milk yield, milk fat, and milk protein production was chosen to test this new approach in a Holstein-Friesian dairy cattle population. The K232A polymorphism and the promoter VNTR (variable number of tandem repeats) of the *DGAT1* gene were used. All lactation curve models predicted the production curves sufficiently. Nevertheless, for predicting genotype effects, the Wilmink curve indicated the closest fit to the data. This study shows that the characteristic gene effects for *DGAT1* genotypes occur after lactation d 40, which might be explained by a link to other genes affecting metabolic traits. Furthermore, allele substitution effects of allele K of the K232A locus showed that the typical effect of low milk and protein yield is due mainly to a lower overall production level, whereas the higher fat and protein content is reached by increased production toward its peak and fat yield is increased because of a higher production after this peak. Predicting gene effects with production curves gives better insight into the timeline of gene effects. This can be used to form genetic groups, in addition to feeding groups, for managing livestock populations in a more effective way.

**Key words:** lactation curve model, gene effect, dairy cow, *DGAT1*

### INTRODUCTION

Statistical modeling of lactation curves based on test-day records is often used to estimate missing 305-d records or in random regression functions to directly estimate breeding values. It has previously been observed that EBV based on test-day records are more accurate than those based on 305-d models. This could be because test-day models take the progress of the whole lactation into account, rather than using a single cumulative measurement or estimate of the lactation yield at d 305 (Sawalha et al., 2005). Estimates for fixed effects are also more precise for the test-day model (Pool and Meuwissen, 1997; Sawalha et al., 2005). Furthermore, estimating parameters for predicting lactation curves from test-day data could provide better insight into the different stages of lactation. Nonetheless, many studies regarding candidate genes for milk composition and quality are carried out using 305-d yield, especially when daughter yield deviations from national breeding evaluation centers are examined. Currently, lactation curve models are mainly used to estimate genetic factors such as heritabilities or biological and environmental effects on the lactation (Druet et al., 2003; Macciotta et al., 2006).

In the last decade, the effects of the K232A locus in exon 8 and a promoter variable number of tandem repeats (**VNTR**) in the acylCoA:diacylglycerol acyltransferase 1 (*DGAT1*) gene on BTA14 have been well described and are shown to be strongly associated with an increase in fat and protein content and decrease in milk yield, making *DGAT1* a strong candidate gene for milk production (Gautier et al., 2007; Kuehn et al., 2007; Näslund et al., 2008; Rahmatalla et al., 2008). The objective of this study was to find an appropriate model for analyzing the gene effects over the period of a lactation cycle and to use this model to demonstrate whether gene and allele effects change over a timeline. Three lactation curve models were used to analyze genotype and allele effects of the 2 loci in the *DGAT1* gene in a German Holstein-Friesian herd. Gene effects were estimated for milk yield, fat and protein yields, and fat and protein contents and compared with a test-day random regression model as a reference. Using these

Received November 23, 2009.

Accepted September 16, 2010.

<sup>1</sup>Corresponding author: gudrun.brockmann@agrar.hu-berlin.de

models, better insight into the *DGAT1* gene effects on the different lactation stages could be gained.

## MATERIALS AND METHODS

### Phenotypic Data

A base population of 1,465 German Holstein-Friesians located on 3 dairy farms in northern Germany with similar management and environment conditions were investigated. Milk samples were collected monthly between February 1999 and July 2006. Milk yield, fat and protein yields, and fat and protein contents were recorded over the first 3 lactations (35,016 records in total). Only cows with a minimum of 8 test-day records per lactation and with at least one record before d 50 of lactation were considered to ensure a high level of accuracy. Test-day records beyond lactation d 340 were not considered (25,895 records after quality control). The cattle population comprised offspring from 262 sires.

### Genotype Data

The *DGAT1* K232A mutation was genotyped according to Winter et al. (2002); the resulting alleles were described as K (Lys) and A (Ala). Genotyping of the VNTR polymorphism was performed according to Kühn et al. (2004), and the alleles were numbered from 1 to 5, where alleles 1, 2, 3, 4, and 5 had 3, 4, 5, 6, and 7 repeats, respectively. Genotype data were taken from Rahmatalla et al. (2008). For the statistical analyses, only genotypes and alleles with a frequency >5% were used, resulting in 1,135 animals for analysis.

### Statistical Analysis

**Lactation Curve Models.** The lactation curves were modeled using 3 approaches that have all been previously used to describe lactation curves for. To date, these lactation curve models have been used mainly to estimate biological and environmental effects or to predict the lactation curve for unfinished lactations and breeding value estimation (Druet et al., 2003; Macciotta et al., 2006; Dematawewa et al., 2007).

**The Wood Model.** The Wood model (**WOOD**) is based on the Wood curve and contains 3 lactation curve parameters. Wood (1967) described his lactation curve function as follows:

$$Y_{\text{DIM}} = a \times \text{DIM}^b \times \exp^{-c \times \text{DIM}}, [\text{WOOD}]$$

where  $Y_{\text{DIM}}$  is the test-day record at a given DIM, parameter  $a$  represents a scaling factor at the begin-

ning of the lactation, and parameters  $b$  and  $c$  describe the incline and the decline of the slope of the curves, respectively.

**The Wilmink Model.** The Wilmink (**WIL**) curve, according to Wilmink (1987), is described as follows:

$$Y_{\text{DIM}} = d + e \times \exp^{-k \times \text{DIM}} + f \times \text{DIM}, [\text{WIL}]$$

where parameter  $d$  represents the level of production, parameter  $e$  the production increase toward peak, and parameter  $f$  the production decrease after the peak. Scale parameter  $k$  was estimated on the whole data set with a nonlinear regression model and fixed for further analyses with  $k = 0.06$  for milk yield,  $k = 0.04$  for fat yield,  $k = 0.05$  for fat content,  $k = 0.01$  for protein yield, and  $k = 0.1$  for protein content.

**The Ali and Schaeffer Model.** The Ali and Schaeffer (**ALI**) model was developed by Ali and Schaeffer (1987) out of 3 lactation curve models:

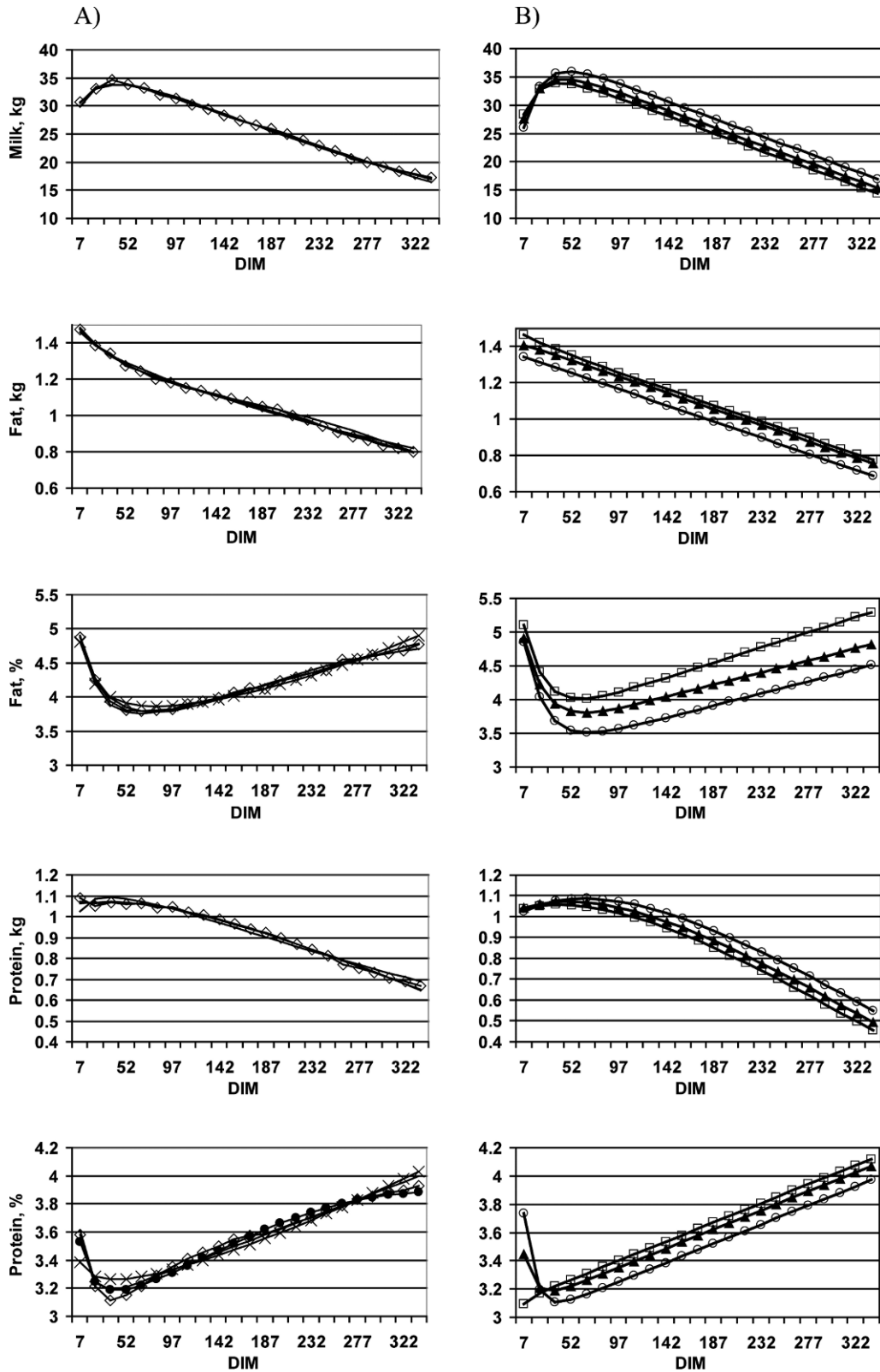
$$Y_{\text{DIM}} = g + h \times \gamma_{\text{DIM}} + i \times \gamma_{\text{DIM}}^2 + j \times R_{\text{DIM}} + l \times R_{\text{DIM}}^2, [\text{ALI}]$$

where  $\gamma_{\text{DIM}} = \text{DIM}/340$ ,  $R_{\text{DIM}} = \ln(340/\text{DIM})$ ; 340 is the number of total lactation days), parameter  $g$  describes the peak yield, parameters  $h$  and  $i$  are associated with the decreasing slope, and parameters  $j$  and  $l$  are associated with the increasing slope.

The NLIN procedure of the SAS program (SAS Institute Inc., Cary, NC) was used to estimate the lactation curve parameters of the WOOD, WIL, and ALI models. For each cow, parameters were estimated for all 3 lactations separately and treated as separate phenotypes for the association study.

**Association Study.** The association between the genotypes and the curve parameters as well as between the milk performance traits in the test-day model was carried out with the MIXED procedure of the SAS program (2008, version 9.2; SAS Institute Inc.).

Days in milk and age at first calving (A) were included in the models as covariates. Calving season (CS), date of the record (D), herd (H), number of lactation (L), and interaction terms  $H \times A$ ,  $H \times D$ ,  $H \times L$ ,  $H \times CS$  were considered as fixed effects. All models and each trait were fitted separately and nonsignificant factors were excluded from the model. Furthermore, the sire and the cow itself were fitted as random effects to account for paternal half-sib relationships and the permanent environmental effects. Calving season was arranged in classes covering the years 1999 to 2006, with each year having 4 quarters.



**Figure 1.** A) Production curves predicted with the Wilkink, Wood, and Ali and Schaeffer models and compared with the real production curve of the population. If curves overlap with the real data curve, they are shown just as a line.  $\diamond$  = real data curve,  $\times$  = Wood curve,  $\bullet$  = Ali and Schaeffer curve. B) Production curves predicted with the Wilkink model for the different genotypes of the *DGAT1* K232A locus: AA ( $\circ$ ), KA ( $\blacktriangle$ ), and KK ( $\square$ ).

**Table 1.** Average parameter estimates per trait for the 3 lactation curve models

Model	Parameter	Milk, kg	Fat, kg	Fat, %	Protein, kg	Protein, %
Wood	<i>a</i>	29.64	1.98	8.28	0.90	3.43
	<i>b</i>	0.15	0.008	-0.14	0.14	-0.01
	<i>c</i>	0.004	0.002	-0.002	0.003	-0.001
Wilmink	<i>d</i>	37.75	1.35	3.41	1.42	3.06
	<i>e</i>	-10.61	0.17	2.22	-0.38	0.14
	<i>f</i>	-0.07	-0.002	0.004	-0.002	0.003
Ali and Schaeffer	<i>g</i>	-44.24	-1.16	8.92	0.11	9.15
	<i>h</i>	103.74	3.64	-6.86	1.49	-7.84
	<i>i</i>	-46.73	-1.85	2.99	-1.07	2.77
	<i>j</i>	50.76	1.40	-3.62	0.60	-3.87
	<i>l</i>	-9.02	-0.20	0.75	-0.10	0.68

A weighting variable  $w = \frac{h^2(n-1)}{4n}$ , where  $n$  = the number of test-day records, to weight the residuals was introduced to account for the different number of test-day records per cow and lactation in the test-day model. The heritability  $h^2$  for the yield traits was set to 0.25, and the  $h^2$  for the content traits was set to 0.5 (an estimated average for the heritabilities was taken from Suzuki and Van Vleck, 1994; Visscher and Goddard, 1995; Al-Seaf et al., 2007; Stoop et al., 2007).

With the lactation curve parameters as dependent variables, model [1] was used for genotype effects. A similar allelic model was used to estimate allele effects, not shown here, because of its overlapping results:

$$y_{ijklm} = \mu + K232A_i + VNTR_j + F_k + S_l + R_m + \varepsilon_{ijklm} \quad [1]$$

where  $y_{ijklm}$  are the lactation curve parameters for each lactation;  $\mu$  is the overall mean;  $K232A_i$  is the fixed effect of the genotype at the K232A locus (AA, KA, KK) and  $VNTR_j$  is fixed effect of the genotype at the VNTR locus (23, 33, 34; genotypes with frequencies >5%);  $F_k$  are the significant fixed effects and interaction terms between the fixed effects as stated above;  $S_l$  is the random sire effect;  $R_m$  are the random cow effects; and  $\varepsilon_{ijklm}$  is the random residual term.

The test-day model was based on Näslund et al. (2008), which used the test-day record directly. Model [2] estimates the average genotype effects on daily milk production, and a similar allelic model was used to estimate the allele effects, respectively:

$$z_{ijklmn} = \mu + K232A_i + VNTR_j + b_{m1}DIM + b_{m2}DIM^2 + b_{m3}\ln(DIM^{-1}) + b_{m4}[\ln(DIM^{-1})]^2 + F_k + S_l + R_m + \varepsilon_{ijklmn} \quad [2]$$

where  $z_{ijklmn}$  is the test-day record,  $\mu$  is the overall mean;  $K232A_i$  is the fixed effect of the genotype at the K232A locus (AA, KA, KK) and  $VNTR_j$  is fixed effect of the genotype at the VNTR locus (23, 33, 34; genotypes with frequencies >5%); DIM is a covariate for the days in milk;  $b_{m1}$  to  $b_{m4}$  are regression coefficients associated with the fixed lactation function;  $F_k$  are the significant fixed effects and interaction terms between the fixed effects as stated above;  $S_l$  is the random sire effect;  $R_m$  are the random cow effects; and  $\varepsilon_{ijklmn}$  is the random residual term.

Lactation curve models were compared regarding their goodness of fit to the mean square error (MSE); the smaller the error variance the better the fit of the model.

## RESULTS

### Goodness of Fit

All lactation curve models led to curves following the shape of the real data with only small inaccuracies, in particular around the peak for highest protein content (Figure 1A). Overall, the differences in MSE were quite small but the WIL model showed the smallest MSE. Comparing the shape of the curves for the different genotypes with each other, the Wilmink curve had the best fit to the real data for milk yield and fat, as well as for protein content. For protein yield, the fit of the Wilmink curve was only the best until d 160, after which the curve started to decrease faster than the real data (Figure 1B).

Estimates of curve parameters under the Wood model produced estimated curves that predicted well above the observed data for all parameters and genotypes (Table 1).

The Ali and Schaeffer curve showed a relatively good fit to the real data with some inaccuracies, particularly at the beginning of the lactation, and a general

**Table 2.** Allele frequencies for the K232A locus and the promoter VNTR (variable number of tandem repeats) in the *DGAT1* gene in a German Holstein-Friesian cow population

Locus	Allele	Frequency
<i>DGAT1</i> K232A	A	0.56
	K	0.44
<i>DGAT1</i> promoter VNTR	1 <sup>1</sup>	0.02
	2	0.18
	3	0.63
	4	0.15
	5 <sup>1</sup>	0.02

<sup>1</sup>Alleles 1 and 5 of the promoter VNTR as well as genotypes containing those alleles were not considered in the statistical analysis due to frequencies <0.05.

underestimation of fat yield for all genotypes (Table 1; Supplementary Figure 1; <http://www.journalofdairyscience.org/>).

### Allele and Genotype Frequencies

Allele frequencies showed a slightly higher proportion of the A allele at the K232A locus and the highest frequency for allele 3 at the promoter VNTR locus followed by a similar proportion of alleles 2 and 4. Alleles 1 and 5 occurred in just 2% of the population (Table 2). The promoter VNTR was not in Hardy-Weinberg equilibrium according to the  $\chi^2$  test (55.99 with 10 df >18.3). Because of the low frequency, alleles 1 and 5 as well as genotypes with a frequency <5% in the population were excluded from further analyses.

### Allele and Genotype Effects

Because of the choice of the WIL curve as the best estimator of production curves, only the effects found for the parameters of the WIL curve are described in more depth.

Allele substitution effects for allele K showed a decrease in the level of the production (parameter *d*) for milk and protein yield and an increase in fat yield and fat and protein contents. The slope toward the peak (parameter *e*) was steeper for the K allele for yield traits. A highly significant ( $P < 0.0001$ ) flattening effect on the slope after the peak (parameter *f*) was observed for fat content. Allele 3 of the promoter VNTR showed a sharpening of the slope before and after the peak for protein content. Allele 4 of the promoter VNTR had a flattening effect on the slope after the peak for fat yield. Other VNTR alleles had no significant effect on either trait (Supplementary Table 1; <http://www.journalofdairyscience.org/>).

All genotypes of the K232A locus showed highly significant effects ( $P < 0.0001$ ) over all traits for all parameters of the WIL model. The highest production level (parameter *d*) was found for genotype AA and the lowest for genotype KK for milk and protein yields. For fat yield as well as fat and protein contents, the genotype KK marked the highest production level. The heterozygous genotype always showed an intermediate production level and a nonsignificant difference between genotype KA and KK for fat and protein yields (Table 3, Figure 1B). For parameter *e*, depicting the slope toward the peak, genotype AA showed the strongest increase for milk and protein yields and the strongest decrease for fat and protein contents. The differences for fat yield and content as well as between genotype KA and KK for all traits except protein content were not significant (Table 3, Figure 1B). The slope after the peak (parameter *f*) showed significant differences between the genotypes for fat content as well as between genotype AA and KK for protein content only (Table 3, Figure 1B). The differences between the VNTR genotypes on any parameter were not significant.

The peak production for milk and protein yields was delayed for about 2 to 3 wk for genotype AA compared

**Table 3.** Least squares means differences between genotypes at the K232A locus in *DGAT1* estimated with model 1 for the parameters of the Wilmink model (LSM  $\pm$  SE)

Parameter	Genotype	Milk, kg	Fat, kg	Fat, %	Protein, kg	Protein, %
<i>d</i>	AA-KA	1.58 $\pm$ 0.39***	-0.07 $\pm$ 0.01***	-0.31 $\pm$ 0.39***	0.06 $\pm$ 0.02*	-0.10 $\pm$ 0.02***
	AA-KK	2.59 $\pm$ 0.48***	-0.09 $\pm$ 0.02***	-0.45 $\pm$ 0.49***	0.10 $\pm$ 0.03**	-0.14 $\pm$ 0.02***
	KA-KK	1.01 $\pm$ 0.41*	-0.02 $\pm$ 0.02†	-0.14 $\pm$ 0.04**	0.04 $\pm$ 0.03†	-0.05 $\pm$ 0.02*
<i>e</i>	AA-KA	-4.92 $\pm$ 1.41**	0.01 $\pm$ 0.04†	0.35 $\pm$ 0.15†	-0.09 $\pm$ 0.03*	0.79 $\pm$ 0.26**
	AA-KK	-7.59 $\pm$ 1.74***	-0.04 $\pm$ 0.05†	0.27 $\pm$ 0.19†	-0.12 $\pm$ 0.04**	1.60 $\pm$ 0.32***
	KA-KK	-2.67 $\pm$ 1.49†	-0.05 $\pm$ 0.05†	-0.08 $\pm$ 0.16†	-0.04 $\pm$ 0.03†	0.81 $\pm$ 0.28**
<i>f</i>	AA-KA	-0.001 $\pm$ 0.002†	0.0001 $\pm$ 0.0001†	-0.00059 $\pm$ 0.0002***	-0.0002 $\pm$ 0.0001†	-0.0002 $\pm$ 0.0001†
	AA-KK	-0.0002 $\pm$ 0.002†	0.0002 $\pm$ 0.0001†	-0.00124 $\pm$ 0.0002***	-0.0002 $\pm$ 0.0001†	-0.0002 $\pm$ 0.0001*
	KA-KK	0.001 $\pm$ 0.002†	0.0001 $\pm$ 0.0001†	-0.00065 $\pm$ 0.0002***	6.62E-6 $\pm$ 0.0001†	-0.0001 $\pm$ 0.0001†

† $P < 0.10$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

**Table 4.** Least squares means differences between genotypes at the K232A locus and the promoter VNTR (variable number of tandem repeats) in *DGAT1* estimated with model 2 (LSM  $\pm$  SE)

Locus	Genotype	Milk, kg	Fat, kg	Fat, %	Protein, kg	Protein, %
K232A	AA-KA	1.29 $\pm$ 0.29***	-0.05 $\pm$ 0.01***	-0.39 $\pm$ 0.03***	0.01 $\pm$ 0.009†	-0.12 $\pm$ 0.01***
	AA-KK	2.45 $\pm$ 0.36***	-0.06 $\pm$ 0.01***	-0.64 $\pm$ 0.04***	0.04 $\pm$ 0.01**	-0.16 $\pm$ 0.02***
	KA-KK	1.16 $\pm$ 0.31**	-0.01 $\pm$ 0.01†	-0.25 $\pm$ 0.03***	0.03 $\pm$ 0.01**	-0.05 $\pm$ 0.01**
VNTR	23-33	-0.29 $\pm$ 0.29†	-0.008 $\pm$ 0.01†	-0.004 $\pm$ 0.03†	-0.006 $\pm$ 0.009†	0.002 $\pm$ 0.01†
	23-34	0.54 $\pm$ 0.33†	0.009 $\pm$ 0.01†	-0.04 $\pm$ 0.03†	0.02 $\pm$ 0.01†	-0.001 $\pm$ 0.02†
	33-34	-0.25 $\pm$ 0.30†	-0.001 $\pm$ 0.01†	0.05 $\pm$ 0.03†	-0.01 $\pm$ 0.01†	-0.001 $\pm$ 0.01†

† $P < 0.10$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

with the other genotypes and reached the maximum between lactation d 50 to 70.

Parameters of the WOOD model showed highly significant effects for all genotypes and traits ( $P < 0.0001$ ) but never for the differences between genotypes of the promoter VNTR. The ALI model presented significant effects on parameters for some traits and for some genotypes only.

Using a test-day model (model [2]), only allele K of the K232A locus showed significant effects ( $P < 0.0001$ ) with a decrease in milk and protein yields and an increase in fat yield and fat and protein contents. All genotypes marking highly significant effects ( $P < 0.0001$ ) indicated a decreasing effect of the KK genotype of the K232A locus compared with the KA genotype and even stronger when compared with the AA genotype for milk and protein yield, whereas an increasing effect was shown for fat yield and fat and protein contents. Allele effects and differences between the promoter VNTR genotypes in the test-day model were not significant (Table 4; Supplementary Table 2, <http://www.journalofdairyscience.org/>).

## DISCUSSION

According to Silvestre et al. (2006), accuracies for the 3 models described, as well as for cubic spline models and Legendre polynomials, do not differ much when based on monthly test-day data. In the present study, enough data points were available to presume sufficient accuracy. To support this, MSE values were very small.

The Wood curve, with only 3 parameters and therefore a rather fixed curve shape, generally over- or underestimated the peak production, also described by Olori et al. (1999). This misprediction is mainly caused by parameters  $b$  (incline toward peak) and  $c$  (decline after peak), which led the curves above the actual production curve. Although lactation curves with 5 parameters were seen to be more flexible and, thus, consistent in their accuracy, a bias was observed in previous studies throughout the lactation for the Ali and Schaefer

curve (Jamrozik and Schaeffer, 1997; Druet et al., 2003). This bias was visible in our study, particularly at the beginning and the end of the lactation. It has been described in some studies that the Wilmink curve, formally a 4-parameter curve, had difficulties in modeling all the variation of the lactation curve because of reduced flexibility (Druet et al., 2003). Nevertheless, it was reported that the Wilmink curve has a satisfactory accuracy and small standard deviation when sufficient data are available (Olori et al., 1999; Macciotta et al., 2006; Silvestre et al., 2006). In our study, the Wilmink curve also seems to be the model of choice due to low residual errors and a good fit in predicting the real production curves. A further discussion will focus on the results of the WIL model.

Genotype and allele frequencies are in concordance with other reports for Holstein-Friesian breeds, and the production curves for all alleles and genotypes of the K232A locus predicted by the WIL model are in agreement with the effects described in our test-day model and other previous studies with a decreasing effect of the K allele on milk and protein yield and an increasing effect on fat yield and fat and protein content (Grisart et al., 2002; Winter et al., 2002; Thaller et al., 2003; Kühn et al., 2004; Gautier et al., 2007; Kuehn et al., 2007; Näslund et al., 2008). Furthermore, the dominance effect of the K allele at the K232A locus could be confirmed, although not specifically tested; genotype differences were smaller between the KK and the KA genotypes than between AA and KA (Kuehn et al., 2007). The difference between the genotypes KK and KA was mostly insignificant for the parameter effects of the WIL model.

Allele substitution effects of allele K at the K232A locus described with the parameters of the WIL model showed that the typical effect of low milk and protein yields is mainly due to a lower overall production level. The higher protein content is reached by increased production toward the peak. Fat content is increased because of a higher production after the peak.

In general, the characteristic genotype effects for the K232A locus occurred after d 40 of the lactation. For

example, animals with the genotype AA, which is usually characterized by a high milk and protein yield and low protein content, showed the lowest milk and protein production combined with the highest protein content until d 30 to 45. This is caused by a low or high production level right at the start of lactation for milk yield or protein content, respectively. Despite an initial lower production of cows with the genotype AA, a steeper slope towards the peak combined with a later/ higher peak causes the overall higher production in comparison to cows of the genotypes KA or KK. Looking at parameter  $f$  of the WIL model (slope after the peak), the insignificant differences between the genotypes imply no difference regarding persistency. Significant differences between the genotypes could be observed for fat content, and indicate better persistency for the KK genotype (Table 3, Figure 1B).

The differences between the effects of the promoter VNTR genotypes did not coincide for all traits for the WIL model and the test-day model but were marginal for both. Only one outstanding difference was observed for protein content, where genotype 33 showed an initial production of about 0.4% higher protein content compared with the other 2 genotypes and the minimum of production occurring 1 wk later. Despite this higher initial production, the test-day model as well as the production curve of the WIL model indicated a negative effect of genotype 33 on all traits, albeit not significant. Allele effects for the promoter VNTR were insignificant for both models.

Effects of VNTR allele 5, as reported by Kuehn et al. (2007) and Gautier et al. (2007), could not be verified due to the absence of this allele in our population.

During the first weeks of lactation, higher producing cows reach their peak production and the high energy demand exceeds the feed energy. Thus, the lactating cows exhibit a negative energy balance, and animals with the *DGAT1* genotype AA appear to compensate best for this negative energy balance. An interaction between the *DGAT1* genotypes and genes involved in metabolism and energy allocation might be an explanation for the reported effects.

Although the content traits have already a high heritability ( $h^2 \sim 0.50$ ), there seems to be more opportunity to influence the yield traits ( $h^2 \sim 0.25$ ) by considering environmental factors. Feeding management changes in most dairy farms from the transition period (2 wk before until 3 wk after calving) to the high and late lactation period around lactation d 30. If feeding groups were built not only regarding the lactation stage but also based on the genotype, it might be possible to enhance the milk production of the less producing genotypes

with further adaptation of the nutrition. Furthermore, modern automated feeding machines could fine-tune the feed supply to the cow's individual requirements.

Nevertheless, using only a single gene seems insufficient. An application of the proposed method on genome-wide association studies looks promising and useful for livestock management and breeding, as well as for enhancing breeding value estimations.

## CONCLUSIONS

The Wilink curve is computationally straightforward with small standard errors; it performs just as well as more complex models in prediction of the gene effects when a sufficient number of data point is available. The estimated genotypic effects of the 2 *DGAT1* polymorphisms coincide with the results of a traditional test-day model and confirm previous studies. However, instead of giving just one value representing the effect of a locus over the whole lactation, the application of lactation curve parameters as phenotypes enables clarification of the actual effect of the candidate gene on different lactation stages. We have shown here that the known genotype effect of the *DGAT1* gene with a low milk yield and high fat content become apparent only after lactation d 40. During the first weeks of the lactation, genotype effects might even be opposite of those observed for later lactation weeks. Furthermore, the lactation stage in which effects are expressed is different for each trait. A possible link between the *DGAT1* gene and genes affecting energy balance was suggested. Analyzing the gene effects of additional candidate genes for milk production traits with lactation curves could provide a new way of enhancing the production level of a herd by providing detailed information on effects during different stages of lactation. This information could be used for a more individual-specific feeding management based on genetic groups in addition to lactation stages. Furthermore, this information could be added to breeding value estimation.

## ACKNOWLEDGMENTS

E. M. Strucken acknowledges the SABRETRAIN project (EC Contract number MEST-CT-2005-020558) funded by Marie Curie Host Fellowship for Early Stage Research Training. S. Rahmatalla acknowledges support from the German Academic Exchange Service (DAAD; Bonn, Germany). D. J. de Koning acknowledges the Biotechnology and Biological Sciences Research Council (BBSRC; Swindon, UK) for contributions through the Institute Strategic Programme Grant.



## REFERENCES

- Al-Seaf, A., J. F. Keown, and L. D. Van Vleck. 2007. Genetic parameters for yield traits of cows treated or not treated with bovine somatotropin. *J. Dairy Sci.* 90:501–506.
- Ali, T. E., and L. R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637–644.
- Dematawewa, C. M. B., R. E. Pearson, and P. M. VanRaden. 2007. Modeling extended lactations of Holsteins. *J. Dairy Sci.* 90:3924–3936.
- Druet, T., F. Jaffrézic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86:2480–2490.
- Gautier, M., A. Capitan, S. Fritz, A. Eggen, D. Boichard, and T. Druet. 2007. Characterization of the *DGAT1* K232A and variable number of tandem repeat polymorphisms in French dairy cattle. *J. Dairy Sci.* 90:2980–2988.
- Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford, P. Berzi, N. Cambisano, M. Mni, S. Reid, P. Simon, R. Spelman, M. Georges, and R. Snell. 2002. Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition. *Genome Res.* 12:222–231.
- Jamrozik, J., and L. R. Schaeffer. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first-lactation Holsteins. *J. Dairy Sci.* 80:762–770.
- Kuehn, C., C. Edel, R. Weikard, and G. Thaller. 2007. Dominance and parent-of-origin effects of coding and non-coding alleles at the acylCoA-diacylglycerol-acyltransferase (*DGAT1*) gene on milk production traits in German Holstein cows. *BMC Genet.* 8:62.
- Kühn, C., G. Thaller, A. Winter, O. R. Bininda-Emonds, B. Kaupe, G. Erhardt, J. Bennewitz, M. Schwerin, and R. Fries. 2004. Evidence for multiple alleles at the *DGAT1* locus better explains a quantitative trait locus with major effect on milk fat content in cattle. *Genetics* 167:1873–1881.
- Macciotta, N. P. P., C. Dimauro, G. Catillo, A. Coletta, and A. Cappio-Borlino. 2006. Short communication: Factors affecting individual lactation curve shape in Italian river buffaloes. *Livest. Sci.* 104:33–37.
- Näslund, J., W. F. Fikse, G. R. Pielberg, and A. Lundén. 2008. Frequency and effect of the bovine acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*) K232A polymorphism in Swedish dairy cattle. *J. Dairy Sci.* 91:2127–2134.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGuirk. 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.* 58:55–63.
- Pool, M. H., and T. H. E. Meuwissen. 1997. Prediction of 305-day milk yield from a limited number of test-days 5 using a test-day model. Pages 59–65 in 16th. *Int. Bull. Eval. Serv.*, Uppsala, Sweden.
- Rahmatalla, S., U. Mueller, E. M. Strucken, and G. Brockmann. 2008. Der Effekt von *DGAT1*-Genvarianten in Deutschen Holstein-Kuehen unter Produktionsbedingungen. *Zuchtungskunde* 80:473–484.
- Sawalha, R. M., J. F. Keown, S. D. Kachman, and L. D. Van Vleck. 2005. Genetic evaluation of dairy cattle with test-day models with autoregressive covariance structures and with a 305-d model. *J. Dairy Sci.* 88:3346–3353.
- Silvestre, A. M., F. Petim-Batista, and J. Colaco. 2006. The accuracy of seven mathematical functions in modelling dairy cattle lactation curves based on test-day records from varying sample schemes. *J. Dairy Sci.* 89:1813–1821.
- Stoop, W. M., H. Bovenhuis, and J. A. M. van Arendonk. 2007. Genetic parameters for milk urea nitrogen in relation to milk production traits. *J. Dairy Sci.* 90:1981–1986.
- Suzuki, M., and L. D. Van Vleck. 1994. Heritability and repeatability for milk production traits of Japanese Holsteins from an animal model. *J. Dairy Sci.* 77:583–588.
- Thaller, G., W. Kramer, A. Winter, B. Kaupe, G. Erhardt, and R. Fries. 2003. Effects of *DGAT1* variants on milk production traits in German cattle breeds. *J. Anim. Sci.* 81:1911–1918.
- Visscher, P. M., and M. E. Goddard. 1995. Genetic parameters for milk yield, survival, workability, and type traits for Australian dairy cattle. *J. Dairy Sci.* 78:205–220.
- Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat, and protein yield for age, season, and stage of lactation. *Livest. Prod. Sci.* 16:335–348.
- Winter, A., W. Kramer, F. A. Werner, S. Kollers, S. Kata, G. Durstewitz, J. Buitkamp, J. E. Womack, G. Thaller, and R. Fries. 2002. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (*DGAT1*) with variation at a quantitative trait locus for milk fat content. *Proc. Natl. Acad. Sci. USA* 99:9300–9305.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165.