Evaluating markers in selected genes for association with functional longevity of dairy cattle

Joanna Szyda1,2*, Małgorzata Morek-Kopeć3, Jolanta Komisarek4, Andrzej Zarnecki5

Abstract

Background: Longevity expressed as the number of days between birth and death is a trait of great importance for both human and animal populations. In our analysis we use dairy cattle to demonstrate how the association of Single Nucleotide Polymorphisms (SNPs) located within selected genes with longevity can be modeled. Such an approach can be extended to any genotyped population with time to endpoint information available. Our study is focused on selected genes in order to answer the question whether genes, known to be involved into the physiological determination of milk production, also influence individual’s survival.

Results: Generally, the highest risk differences among animals with different genotypes are observed for polymorphisms located within the leptin gene. The polymorphism with a highest effect on functional longevity is LEP-R25C, for which the relative risk of culling for cows with genotype CC is 3.14 times higher than for the heterozygous animals. Apart from LEP-R25C, also FF homozygotes at the LEP-Y7F substitution attribute 3.64 times higher risk of culling than the YY homozygotes and W homozygotes at LEP-A80V have 1.83 times higher risk of culling than AA homozygotes. Differences in risks between genotypes of polymorphisms within the other genes (the butyrophilin subfamily 1 member A1 gene, BTN1A1; the acyl-CoA:diacylglycerol acyltransferase 1 gene, DGAT1; the leptin receptor gene, LEPR; the ATP-binding cassette sub-family G member 2, ABCG2) are much smaller.

Conclusions: Our results indicate association between LEP and longevity and are very well supported by results of other studies related to dairy cattle. In view of the growing importance of functional traits in dairy cattle, LEP polymorphisms should be considered as markers supporting selection decisions. Furthermore, since the relationship between both LEP polymorphism and its protein product with longevity in humans is well documented, with our result we were able to demonstrate that livestock with its detailed records of family structure, genetic, and environmental factors as well as extensive trait recording can be a good model organism for research aspects related to humans.

Background

The major goal of our study was to evaluate the association between SNP markers located within selected genes and functional longevity of dairy cows. For this purpose effects of SNPs within genes representing different groups regarding their effect on milk production were selected: (i) genes with well characterised high effects on milk production (two genes), (ii) genes which effect on milk production is contradictory, depending on the analysed data set (two genes), (iii) genes which effects on milk production have not been reported (one gene).

Longevity expressed as the number of days between birth and death is a trait of great importance for both human and animal populations. In our analysis we use dairy cattle as an organism for which we demonstrate how the influence of particular genes on longevity can be modelled. Such an approach can be extended to any genotyped population with time to endpoint information available. On one hand dairy cattle makes a good model organism since it has precise current and historical pedigree data as well as information on herd management, various traits, medical examinations etc. recorded.
through individuals lifetime. On the other hand however, the peculiarity of the survival analysis of dairy cattle data is that in this case the endpoint (i.e. culling) is not only dependent on individuals’ health status, but involves many factors like individual’s production level or economic issues. Consequently, in dairy cattle breeding one can differentiate between true longevity - where the actual probability of animal’s survival is considered and functional longevity - where additional effects expressing the risk of culling a cow irrespective of her production performance level and herd management are included into the modelling of the probability of survival [1].

Functional longevity is rapidly gaining importance as a selection criterion in dairy cattle. Therefore many studies have analyzed its relationship with other characters routinely recorded in dairy cattle, such as reproduction traits [2], udder health traits [3], and level of inbreeding [4] etc. However, from a geneticist’s perspective knowledge whether there exist genes which influence functional longevity is of primary importance.

The genetic effect on survival has been considered in various aspects by other authors. Probably the most widespread approach is to model a cumulated additive effect of all genes, called a polygenic effect, which is routinely estimated for sires of cows from production populations in many countries [5]. Diao et al. [6], Diao and Lin [7], Moreno et al. [8] and Johannes [9] estimated position and effect of quantitative trait loci using linkage analysis applied to simulated and real data. Tregouet et al. [10] and Souverein et al. [11] estimated SNP haplotype effects on failure time for humans. Our study is focused on particular candidate genes in order to answer the question whether selected genes, known to be involved into the physiological determination of milk production, also influence individual’s survival.

**Results**

All comparisons of the most parsimonious model M1 with nine other models including a single polymorphism (M2-M10) as well as comparisons of the full model M20 with nine other models including a single polymorphism resulted in the improvement of fit (Table 1; Figure 1A, 1B). On the other hand differences arise between a functional and a true longevity model. As far as true longevity is concerned, the most significant polymorphism with the largest effect on survival (Table 1; Figure 1C). Using either an unspecified baseline hazard function underlying the Cox model, or a specified Weibull baseline hazard function, did not affect the estimates of polymorphisms markedly. LEP-R25C was the polymorphism with the largest effect on survival (Table 1; Figure 1A, 1B). On the other hand differences arise between a functional and a true longevity model. As far as true longevity is concerned, the most significant polymorphism is DGAT1-K232A (FDR = 0.1341, nominal P = 0.0149) with the highest risk of culling corresponding to heterozygous cows (Figure 1C).

Other (than SNP) effects with significant impacts on functional survival were estimated for protein yield level, lactation number × stage of lactation, and year × season. For those effects risk of culling in first lactation was higher than in subsequent lactations, within lactation always most prone to culling were cow in last stage, high risk was associated with low protein production level. True longevity was also significantly affected by...
lactation number × stage of lactation, and year × season effects.

**Discussion**

The advantage of using a survival analysis model, which takes into account data censoring, was demonstrated by Moreno *et al.* [8]. Using simulated data, the authors showed that for data sets containing censored observations, power of quantitative trait loci detection and accuracy of the estimation of its position and effect is higher for survival models with Weibull and Cox hazard functions than for a standard QTL analysis model, which
ignores censoring. The adequacy of modeling gene effects on survival data was also shown in the analysis of real data by Diao et al. [6]. Performing survival linkage analysis for time to death after bacterial infection of mice using a Weibull hazard function the authors were able to confirm previously identified quantitative trait loci as well as to found a novel locus.

Our study concerned the association of functional longevity with SNPs located within selected genes and found some, albeit not strong, effect of LEP. Although there exist studies which have found significant association of polymorphisms identified within the leptin gene with production traits (e.g. [12]), relation between LEP polymorphism and fitness, metabolism or health related traits, which may be components of functional longevity, are much less common and their results are contradictory. Banos et al. [12] failed to observe a significant association of LEP with body energy traits, but Liefers et al. [13] and Nkrumah et al. [14] reported significant effects of LEP on growth rate, live weight, metabolic body weight, feed intake, and feeding duration. Chebel et al. [15] observed significant associations of LEP with disease risk. Oikonomou et al. [16] detected significant effects of LEP on body energy and blood metabolic traits. Brickell et al. [17] reported significant associations between polymorphisms in the leptin gene with stillbirths and calf mortality. Anyhow, the involvement of leptin gene in inflammation and immune responses [18] makes it a good candidate for a genetic component of longevity.

Although it could have been hypothesized that genes with high effect on yield may also influence longevity, which could be due to redirecting body energy resources towards milk, fat or protein production, no evidence for such phenomenon was observed in our study. The effects of polymorphism located within genes of high impact on milk performance (DGAT1, ABCG2) were not different from a gene with no effect documented (BTN1A1).

Conclusions
As it could have been expected, when information on cows’ production level is ignored in the model corresponding to a true longevity, the most significant effect is attributed to DGAT1 - a polymorphism well known to be responsible for milk production traits in dairy cattle [19]. Cows with a low production level are then culled earlier than their high producing herdmates.

Our results indicate an association between LEP and functional longevity. As such functional longevity is a complex trait, involving various physiological components acting throughout a cow’s production life. On one hand our results are very well supported by other studies involving traits like health, feeding behavior, or reproduction, which are components of functional longevity. On the other hand, the reliability of results obtained is hampered by a moderate size of the data set. Still, in view of the growing importance of functional traits in dairy cattle, LEP polymorphisms should be considered as markers supporting selection decisions.

The relationship between both LEP polymorphism and its protein product with longevity in humans has been documented [20,21] indicating the impact of the gene on reduction in visceral fat, blood glucose, and insulin levels together with increased concentration of ghrelin [22]. With our results we were able to demonstrate that livestock with its detailed records of family structure, genetic, and environmental factors as well as extensive trait recording can be a good model organism for research aspects related to humans.

Methods
DNA samples
DNA samples were extracted from 566 Polish Holstein-Friesian cows for which genotypes at 9 functional single nucleotide polymorphisms (SNP) located within 5 genes were identified:

- a) the butyrophilin subfamily 1 member A1 gene (BTN1A1, GeneID:282157) located on BTA23: P35Q [23] and K468R [24];
- b) the acyl-CoA:diacylglycerol acyltransferase 1 gene (DGAT1, GeneID:282609) on BTA14: K232A [19; 25];
- c) the leptin receptor gene (LEPR, GeneID:497205) on BTA3: T945M [26];
- d) the leptin gene (LEP, GeneID:280836) on BTA4: Y7F [27], R25C and A80V [28], a C/T substitution at position -963 [29];
- e) the ATP-binding cassette sub-family G member 2 gene (ABCG2, GeneID:536203) on BTA6: Y581S [30].

Regarding the effects of the selected genes on milk production traits - i.e. traits posing the major artificial selection pressure in dairy cattle populations, DGAT1 and ABCG2 are known to affect milk production traits, for LEP and LEPR contradictory results have been obtained, while the effect of BTN1A1 has not been reported. Genotypes were determined using PCR-RFLP method, as described by Syzda and Komisarek [31] for BTN1A1, DGAT1, LEP, LEPR and by Komisarek and Dorynek [32] for ABCG2.

Data structure
The analyzed data consisted of 567 genotyped cows, daughters of 109 sires. No missing genotypes were present in the analysed sample, but for ABCG2-Y581S and Y2F, with 0 and 1 cows respectively, the genotype class counts were strongly skewed Table 2. Minor allele frequency varied between 0.013 for ABCG2-Y581S and 0.481 for LEP-R25C. The average number of daughters
per sire was 5 and ranged from 1 to 67. The cows had their production records in four herds. The distribution of cows among herds was not uniform, with 80% of cows (453) active in the same herd. 65% (368) of the cows were represented by individuals with observed culling dates (i.e. uncensored records), with the average failure time of 1 173 days, varying between 32 and 3 168 days. The remainder of the cows comprised animals with so called censored records, represented by cows which have not been culled within the data collection timeframe. For them the censoring times varied between 292nd and 3 046th day, with the average censoring time amounting to 1489 days. For this data the following classes, corresponding to the Polish national routine genetic evaluation model for functional longevity were considered: year × season - comprising years 1999 to 2009 and 2 seasons: April - September and October - March; lactation number × stage of lactation - comprising the first 5 and pooled later lactations and 4 stages of lactation (1-29, 30-179, 180-304, and >304 day of lactation); relative change of herd size from the current year to the next year at April 1st (<-50%, -50% to -30%, -30% to -10%, -10% to 10%, 10% to 30%, 30% to 50%, > 50%); classes of 305-day fat and protein yield levels relative to herd means, defined separately for the first and later lactations (<-50%, -50% to -40%, -40% to -30%, -30% to -20%, -20% to -10%, -10% to 0%, 0% to 10%, 10% to 20%, 20% to 30%, 30% to 40%, 40% to 50%, > 50%); monthly classes of age at first calving (<20, 21, 22,…, >40 month); 3 classes of SNP genotypes represented by two homozygotes and a heterozygote.

Statistical model
The analysed trait is defined as the number of days between the first calving and culling (uncensored records) or the last test day (censored records). The following sequence of functional longevity survival models based on Weibull hazard function was applied to the data:

\[ M_1: \]
\[ h(t) = h_0(t) \exp[ys(t) + sl(t) + hsize(t) + fat(t) + prot(t) + age] \]

\[ M_{11}-M_{19}: \]
\[ h(t) = h_0(t) \exp[ys(t) + sl(t) + hsize(t) + fat(t) + prot(t) + age + SNP_i] \]

\[ M_{20}: \]
\[ h(t) = h_0(t) \exp[ys(t) + sl(t) + hsize(t) + fat(t) + prot(t) + age + P35Q + K468R + K232A + T945M + Y7F + R25C + A80V + C/T + Y581S] \]

where \( h_0(t) = \rho(t)^{-\lambda} \exp[\rho \log(\lambda)] \) represents a baseline Weibull hazard function with scale parameter \( \lambda \) and shape parameter \( \rho \), \( ys(t) \) is a time-dependent fixed effect of year-season, \( sl(t) \) is a time-dependent effect of lactation number × stage of lactation, \( hsize(t) \) is a time-dependent fixed effect of yearly herd size variation, \( fat(t) \) and \( prot(t) \) are time-dependent fixed effects of within herd-year-season classes of 305-day fat and protein production level, \( age \) is a time-independent fixed effect of age at first calving, \( SNP_i \) represents a time independent additive effect of a single polymorphism \( i \) \( (i \in \{P35Q, K468R, K232A, T945M, Y7F, R25C, A80V, \) C/T substitution at position -963, and Y581S\}), \( SNP_{-1} \) represents a time independent additive effect of eight of the polymorphisms excluding \( i \). For comparison, the effects underlying \( M_{20} \) were also estimated using a Cox model in which \( h_0(t) \) is not expressed by a predefined function what corresponds to making no assumption on the underlying baseline risk. Note, that all the above models represent functional longevity models, which aim to estimate effects of SNPs on culling risk irrespectively of cows’ productive performance. Since only animals not culled for milk production contribute meaningful information on longevity, the effects of production level (fat and protein yield class effects) were included into the model in order to make estimates of SNP effects independent from production level. Additionally a modification of \( M_{20} \) with effects of fat and protein production level removed, corresponding to a true longevity was applied to the data. The estimation of model parameters was performed using Survival Kit Version 3.12 [33].

Hypotheses testing
The hypothesis of interest was whether some of the considered polymorphisms influence the risk of cow’s culling and was tested using the likelihood ratio test: \( \lambda = -2[\ln L(\hat{\beta}_0) - \ln L(\hat{\beta}_1)] \), where \( L(\hat{\beta}_0) \) and \( L(\hat{\beta}_1) \)
represent maximum of likelihood functions obtained under the more parsimonious and the less parsimonious model, respectively. In this analysis model parsimony is expressed by the number of polymorphisms considered, while the other model parameters remain the same for all the models. Asymptotically, \( \lambda \) follows the \( \chi^2 \) distribution with the degrees of freedom equal to the difference in the number of effects in the compared models.

All experiments were carried out in compliance with ethical requirements and were approved by the Local Ethics Committee for Animal Research (permission No. 25/2003).

Acknowledgements
Authors thank the Polish Federation of Cattle Breeders and Dairy Farmers for providing the longevity data. The research was supported by the State Committee for Scientific Research projects grant no N N311 450434.

Author details
1Department of Animal Genetics, Wrocław University of Environmental and Life Sciences, Koźlechowska 7, 51-631 Wrocław, Poland. 2Institute of Natural Sciences, Wrocław University of Environmental and Life Sciences, Nowodworska 25, 50-375 Wrocław, Poland. 3Department of Genetics and Animal Breeding, University of Agriculture in Cracow, Mickiewicza 24/28, 30-059 Kraków, Poland. 4Department of Cattle Breeding and Milk Production, Poznan University of Life Sciences, Wojska Polskiego 71A, 60-625 Poznań, Poland. 5National Research Institute of Animal Production, Balice, 32-083 Kraków, Poland.

Authors' contributions
JS provided the idea for the study, as involved in statistical analysis of the data and wrote the manuscript. MMK edited data and performed all computations. JK collected DNA samples and provided genotypes. AZ provided intellectual support and helped in manuscript preparation. All authors read and approved the manuscript.

Received: 28 October 2010 Accepted: 10 March 2011
Published: 10 March 2011

References
1. Ducrocq V, Quasas RL, Pollak E, Casella G. Length of productive life of dairy cows. I Justification of a Weibull model. J Dairy Sci 1988, 71:3061-3070.
2. Sewalem A, Miglor F, Kistemaker GJ, Sullivan P, VanDoormaal BJ. Relationship between reproduction traits and functional longevity in Canadian Dairy Cattle. J Dairy Sci 2008, 91:1660-1668.
3. Neerhof HJ, Madsen P, Ducrocq VP, Völlema AR, Jensen J, Korsgaard IR. Relationship between mastitis and functional longevity in Danish Black and White dairy cattle estimated using survival analysis. J Dairy Sci 2000, 83:1064-1071.
4. Sewalem A, Kistemaker GJ, Miglor F, VanDoormaal BJ. Analysis of inbreeding and its relationship with functional longevity in Canadian Dairy Cattle. J Dairy Sci 2006, 89:2210-2216.
5. Ducrocq V. Two years of experience with the French genetic evaluation of dairy bulls on production-adjusted longevity of their daughters. Interbull Bulletin 1999, 21:60-67.
6. Diao G, Lin DY, Zou F. Mapping quantitative trait loci with censored observations. Genetics 2004, 168:1689-1698.
7. Diao G, Lin DY. Semiparametric methods for mapping quantitative trait loci with censored data. Biometrics 2005, 61:789-798.
8. Moreno CR, Eisen JM, Le Roy P, Ducrocq V. Interval mapping methods for detecting QTL affecting survival and time-to-event phenotypes. Genet Res 2005, 85:139-149.
9. Johannes F. Mapping Temporally Varying Quantitative Trait Loci in Time-to-Failure Experiments. Genetics 2007, 175:855-865.
10. Tregouet DA, Tiert L. Cox proportional hazards survival regression in haplotype-based association analysis using the Stochastic-EM algorithm. Eur J Hum Genet 2004, 12:971-974.
11. Sauveteur GW, Zweitdeman AH, Jukema JW. Time-MTt: Estimating effects of rare haplotypes on failure time using a penalized Cox proportional hazards regression model. BMC Genetics 2008, 9:9.
12. Banos G, Woolliams JA, Woodward BW, Forbes AB, Coffey MP. Impact of Single Nucleotide Polymorphisms in leptin, leptin receptor, growth hormone receptor, and DGAT1 gene loci on milk production, feed and body energy traits of UK dairy cows. J Dairy Sci 2008, 91:3190-3200.
13. Liefers SC, tePas MF, Veerkamp RF, van der Lende T. Associations between leptin gene polymorphisms and production, live weight, energy balance, feed intake, and fertility in Holstein heifers. J Dairy Sci 2002, 85:1633-1638.
14. Nkumah DJ, Li C, Yu J, Hansen C, Keisler DH, Moore SS. Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feed intake, feeding behavior, and measures of carcass merit. J Anim Sci 2005, 83:20-28.
15. Chebel RC, Susca F, Santos JE. Leptin genotype is associated with lactation performance and health of Holstein cows. J Dairy Sci 2008, 91:2893-2900.
16. Oikonomou G, Angelopoulou K, Arsenos G, Zygoianni D, Banos G. The effects of polymorphisms in the DGAT1, leptin and growth hormone receptor gene loci on body energy, blood metabolical and reproductive traits of Holstein cows. Anim Genet 2009, 40:10-17.
17. Brickell JS, Pollott GE, Clemson AM, Otter N, Watkins DC. Polymorphisms in the bovine leptin gene associated with perinatal mortality in Holstein-Friesian heifers. J Dairy Sci 2010, 93:340-347.
18. Wylie ARG. Leptin in farm animals: where are we and where can we go? Animal 2010, 5:246-267.
19. Giurti I, Coppieters W, Famir F, Kariim L, Ford C, Berzi P, Cambisano N, Mni M, Reid S, Simon P, Spelman R, Georges M, Snell R. Positional candidate cloning of a QTL in dairy cattle: Identification of a missence mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Res 2002, 12:223-231.
20. Khabour OF, Mesmar FS, Alatoum MA, Gharaiab KH. Associations of polymorphisms in adiponectin and leptin genes with men’s longevity. Aging Male 2010, 13:188-193.
21. Sloboda DM, Beedle AS, Cupido CL, Gluckman PD, Vickers MH. Impaired Perinatal Growth and Longevity: A Life History Perspective. Curr Gerontol Geriatr Res 2009, Article ID 685740.
22. Boghossiana S, Uenoa N, Dubeb MG, Kalrab P, Kalraa S, Kaira S. Leptin gene transfer in the hypothalamus enhances longevity in adult monogenic mutant mice in the absence of circulating leptin? Neurobiol Aging 2007, 28:1594-1604.
23. Seyfert HM, Lutthen F. The structure of the bovine butyrophilin encoding gene differs grossly from mouse concerning promoter localization and exon organization of the 5-untranslated region. 6th World Congress of Genetics Applied to Livestock Production 1998, 25:51-54.
24. Taylor C, Everest M, Smith C. Restriction fragment length polymorphism in amplification products of the bovine butyrophilin gene: Assignment of bovine butyrophilin to bovine chromosome 23. Anim Genet 1996, 27:183-185.
25. Winter A, Krämer W, Wiemer FAO, Kollers S, Kata S, Durstewitz G, Winter A, Wiemer FAO, Kollers S, Kata S, Durstewitz G. The exon organization of the S-untranslated region. J Dairy Sci 2002, 85:3190-3200.
26. Liefers SC, Veerkamp RF, tePas MF, Delavaud C, Chilliard Y, van der Lende T. Missense mutation in the bovine leptin receptor gene is associated with leptin concentrations during late pregnancy. Anim Genet 2004, 35:138-141.
27. Lagongag R, Wiemer P, Pilla F, Woolliams JA, Williams J. A new mutation in the coding region of the bovine leptin gene associated with feed intake. Anim Genet 2003, 34:371-374.
28. Konfortov BA, Licence VE, Miller JR. Re-sequencing of DNA from a diverse population of bovine butyrophilin to bovine chromosome 23. Anim Genet 1996, 27:183-185.
30. Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Everts-van der Wind A, Lee JH, Drackley JK, Band MR, Hernandez AG, Shani M, Lewin HA, Weller JI, Ron M. Identification of a missense mutation in the bovine ABCG2 gene with major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res* 2005, 15:936-944.

31. Szyda J, Komisarek J. Statistical modeling of candidate gene effects on milk production traits in dairy cattle. *J Dairy Sci* 2007, 90:2971-2979.

32. Komisarek J, Dorynek Z. Effect of ABCG2, PPARGC1A, OLR1 and SCD1 gene polymorphism on estimated breeding values for functional and production traits in Polish Holstein-Friesian bulls. *J Appl Genet* 2009, 50:125-132.

33. Ducrocq V, Sölkner J. The Survival Kit - V3.0, a package for large analyses of survival data. 6th World Congress of Genetics Applied to Livestock Production 1998, 27:447-448.

doi:10.1186/1471-2156-12-30

Cite this article as: Szyda et al. Evaluating markers in selected genes for association with functional longevity of dairy cattle. *BMC Genetics* 2011 12:30.