

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Short communication:

A single nucleotide polymorphism in the bovine *beta-casein* promoter region across different bovine breeds.

A.F. Keating^{*}, T.J. Smith[°], R.P. Ross[†] and M.T. Cairns^{*}

^{*}National Diagnostics Centre, National University of Ireland, Galway, Ireland.

[°] National Centre for Biomedical Engineering Science, National University of Ireland, Galway.

[†] Teagasc, Dairy Products Research Centre, Moorepark, Fermoy, Co. Cork, Ireland.¹

¹Corresponding Author: Dr. Michael Cairns, National Diagnostics Centre, National University of Ireland, Galway. Email: Michael.cairns@nuigalway.ie.

The nucleotide sequence data reported in this paper have been submitted to GenBank and has been assigned the accession number AJ973327.

29 **Introduction:**

30 The bovine *beta-casein* (*CSN2*) gene has been shown to span a region of 8.5kb,
31 containing nine exons and eight intervening introns (Bonsing *et al.*, 1988; Martin *et*
32 *al.*, 2002). The exons range in size from 24 to 498bp, however the introns are much
33 larger and account for 85% of the gene. Twelve genetic variants in the coding
34 sequence of the *beta-casein* gene have been reported (Farrell *et al.*, 2004). The A²
35 allele of the *beta-casein* gene has been associated with a higher milk production (Lin
36 *et al.*, 1986; Bech and Kristiansen, 1990) while the B variant has been associated with
37 an increase in protein content and better cheese-making properties (Marziali and Ng-
38 Hang-Kwai, 1986). The *beta-casein* gene codes for a protein of 209 amino acids with
39 varying regions at codons 67, 106 and 122. The A¹ variant differs from A² at position
40 67, where a histidine replaces a proline (Lien *et al.*, 1992). The *beta-casein* A² variant
41 has histidine and the A³ variant has glycine at position 106 (Lien *et al.*, 1992); the
42 *beta-casein* A² variant has serine at position 122 and the *beta-casein* B variant has
43 arginine at this codon (Stewart *et al.*, 1987; Damiani *et al.*, 1992).

44

45 The *beta-casein* promoter has been characterised and contains a number of binding
46 sites for transcription factors c/ebp, Stat5, Oct and GR (Doppler *et al.*, 1995; Lechner
47 *et al.*, 1997; Raught *et al.*, 1995). A *beta-casein* enhancer element, sited in the distal
48 bovine promoter between -1562 to -1613, contains binding sites for Stat5, c/ebp,
49 YY1 and GR (Raught *et al.*, 1995). In addition, analysis of the murine *beta-casein*
50 promoter has shown the functional significance of the Runx2 transcription factor in
51 full transcriptional activation of the *beta-casein* gene (Inman *et al.*, 2004).
52 Polymorphisms have been investigated in the *beta-casein* gene promoter of different
53 bovine breeds. The *beta-casein* promoters from Jersey, Brown Swiss and Holstein

54 bulls (one of each), were sequenced and the only difference found was a single base
55 deletion at position -516 (Bleck *et al.*, 1996). Four additional sequence differences
56 (single base deletions) were found when comparing sequences to the database
57 sequence, however these are more likely to be sequencing errors in the original
58 sequence (Bonsing *et al.*, 1988). Another investigation into the incidence of
59 polymorphic sites in the *beta-casein* gene promoter identified seven polymorphic sites
60 in the region (Schild and Geldermann, 1996). A study by Szymanowska *et al.* (2004)
61 screened Polish Black-and-White (n = 81) and Polish Red (n = 195) cows for the
62 incidence of the G to C change at -109 identified in the Schild and Geldermann study
63 (1996) but no polymorphism was identified (Szymanowska *et al.*, 2004). Promoter
64 studies have not indicated that differences in casein gene expression are due to these
65 variations but it has been suggested that gene expression changes may instead result
66 from a combination of promoter variants, i.e. that certain haplotypes influence casein
67 gene expression (Martin *et al.*, 2002).

68

69 In this study, polymorphism incidence in the *beta-casein* gene promoter in nine
70 bovine breeds typical of the Irish herd were investigated. The bovine breeds chosen
71 included dairy, dual-purpose and non-dairy (beef) breeds. Potential links between
72 promoter polymorphisms and structural gene polymorphisms were also investigated.

73 **Materials and Methods:**

74 *DNA isolation.* Blood was obtained from the coccygeal vein of animals from nine
75 bovine breeds, namely high genetic merit Holstein-Friesian (n=4), low genetic merit
76 Holstein-Friesian (n=4), Irish-Friesian (n=4), Dutch-Friesian (n=4), Limousin (n=6),
77 Montbeliarde (n=4), Charlois (n=2), Normande (n=4), Norwegian Red (n=2) and
78 Kerry (n=8). DNA extractions were carried out using the Gentra Capture Column™
79 (Gentra, UK) system from approximately 200µl of whole blood per animal. Blood
80 was stored at –80°C and DNA was stored at –20°C until further use.

81 *Polymerase Chain Reaction.* Primers (located at positions 97-120 and 1799-1824 in
82 the NCBI database sequence X14711) were designed to amplify a 1728bp fragment of
83 the *β-casein* gene promoter (MWG Biotech, UK). A second set of primers (located at
84 positions 7574 – 7593 and 8287 – 8306 in NCBI database sequence M55158) were
85 used to amplify a 732bp fragment encompassing the polymorphism that distinguishes
86 the A¹ and A² coding sequence variants. PCR was carried out from a starting
87 template of approximately 200ng of genomic DNA in a final volume of 50µl
88 containing 1X *Taq* DNA polymerase buffer (Invitrogen, UK), 1.5mM MgCl₂, 200µM
89 dNTPs (Promega, UK), 0.3µM each primer, and 1U *Taq* polymerase (Invitrogen,
90 UK). Conditions were an initial incubation at 95°C for 2min, followed by 35 cycles
91 of 95°C for 1min, 58°C for 1min and 72°C for 1min.

92 *Restriction digestion.* Digestion of PCR products was carried out in a final volume of
93 20µl containing 10µl of PCR product, 1X reaction buffer and 1U *EcoRI* restriction
94 enzyme. Reactions were incubated at 37°C for 2h and resolved on a 2% agarose gel
95 in 1X Tris Borate EDTA (TBE) buffer at 90V for 1h.

96 *Sequencing and bioinformatics.* Sequencing of PCR products was carried out by
97 MWG Biotech (Germany). The resulting sequences were analysed using the Vector

98 NTI® Suite of software (Informax™, US). Alignment of sequences for all 42 animals
99 was carried out, and potential polymorphic sites identified. Examination of
100 chromatogram sequence files to detect homozygotic and heterozygotic animals was
101 also performed.

102 *Statistical analysis.* Observed allele frequencies were analysed for equilibrium using
103 the Hardy Weinberg equation. Results were analysed by chi-square test to determine
104 whether observed allele frequencies and allele frequencies predicted by the Hardy
105 Weinberg equations were significantly different. Results of promoter and coding
106 sequence variant screens were analysed by chi-square test. Null hypothesis was that
107 no association occurred between variants.

108

109

110 **Results and Discussion:**

111

112 *Bovine breeds chosen.* Blood samples were obtained from nine bovine breeds chosen
113 to represent the animals typical of the Irish herd, but also to increase the likelihood of
114 genetic variation. The breeds chosen were: dairy - high and low genetic merit
115 Holstein Friesian (n=4 of each), Irish Friesian (n=4) and Dutch Friesian (n=4); dual-
116 purpose - Norwegian Red (n=2), Normande (n=4), Montbeliarde (n=4) and Kerry
117 (n=8); and beef - Limousin (n=6) and Charlois (n=2).

118

119 A single nucleotide polymorphism (SNP) from T to A was identified at position -851
120 from the transcriptional start site. In addition, it was noted that compared to the
121 database sequence all animals had a T insertion at -848. These two variations
122 introduced a recognition site for the *EcoRI* restriction enzyme that allowed
123 development of a PCR-RFLP rapid screen to determine which variant of the β -casein
124 promoter is present (either β -TT, β -TA or β -AA) (Figure 1). The TT allele was
125 undigested and showed a band of 1728bp. The AA allele was digested through the
126 introduction of an *EcoRI* site and showed two bands of 880bp and 848bp. The
127 heterozygote TA allele showed bands at 1728bp, 880bp and 848bp. The incidence of
128 the T/A SNP was TT-45%, TA-38% and AA-17%. The allele frequencies are in
129 Hardy Weinberg equilibrium ($p = 0.97$). When breed differences were observed the
130 incidence of the A allele differed between breeds. The dairy breeds had an incidence
131 of 50%, compared with dual-purpose breeds with an incidence of 38% and the beef
132 breeds with a 100% incidence (Table 1). No transcription factor has as yet been
133 identified that binds at this location, however the study by Schild and Geldermann
134 (1996) suggests that the progesterone receptor may bind at this location.

135

136 The *β-casein* exon VII was also analysed for the presence or absence of the base
137 change at position 67 which encodes either the A¹ or A² variants. This base change is
138 also present in the B variant, so for the purposes of this study that A¹ + B are
139 designated A¹. The promoter and coding sequence genetic variants for the *β-casein*
140 gene for all forty-two animals screened are listed in Table 2. These allele frequencies
141 were also in Hardy Weinberg equilibrium (p = 0.24). The results of the promoter and
142 coding sequence variant screen were analysed statistically and it was indicated that an
143 association existed between the coding sequence variant A²A² and the promoter β-AA
144 variant pair and also between the coding sequence variant A¹A¹ and the promoter
145 variant β-TT (p = 0.00002154). Analysis of a larger group of animals is required to
146 confirm these findings.

147

148 The occurrence of polymorphism in the *β-casein* gene promoter may have an effect
149 on the transcriptional activity of the gene and thus provide an opportunity to improve
150 expression of this important milk protein gene. A previous study of polymorphisms in
151 the promoter region examined thirteen animals and noted seven potential sites of
152 variability (Schild and Geldermann, 1996). However, five of these polymorphisms
153 were seen in only one of the fourteen animals analysed (a different animal with each
154 polymorphism). A further study to determine the incidence of the C to G change at –
155 109 identified in this original screen did not show the change in any of a large number
156 of animals (n = 276) (Szymanowska *et al.*, 2004). Although two of these seven
157 polymorphisms would appear to be quite common, the other five may be rare and only
158 found in specific breeds.

159 In this present study, ten animals were originally screened for polymorphism in the
160 entire 1728bp region of the β -casein promoter. In all ten animals sequenced, a T
161 insertion appeared at -848 which differs from the original database sequence
162 (Bonsing *et al.*, 1998). This insertion was also noted in other studies suggesting that
163 the original sequence is incorrect at this position. (Schild and Geldermann, 1996;
164 Bleck *et al.*, 1996). The only other variable site noted was a T to A base change at -
165 851. This change was noted in four of the ten animals sequenced and fortuitously
166 introduced a recognition site for the *EcoRI* restriction enzyme. This change was also
167 noted in the Schild and Geldermann (1996) study. Forty-two animals were screened
168 by PCR-RFLP and although the number of animals screened per breed was small, the
169 differences between animals bred for different production purposes was noteworthy.
170 In dairy animals the A allele frequency was 50%, with a homozygous AA genotype
171 frequency of 5.5%. In beef animals, however, the A allele frequency was 100%, with
172 a homozygous AA genotype frequency of 25%. The animals bred for both beef and
173 dairy (dual-purpose) were also higher than dairy animals with an A allele frequency of
174 38% and a homozygous AA frequency of 25%.

175

176 **Acknowledgements:**

177 This work was supported by Enterprise Ireland and a Teagasc Walsh fellowship to
178 A.F. Keating.

179

180 **References:**

181 Bech, A.M. and Kristiansen, K.R. (1990). "Milk protein polymorphisms in Danish
182 Dairy cattle and influence of genetic variants on milk yield." J. Dairy Res. **57**: 53-62.

183

184 Bleck, G. T., J.C. Conroy and M.B. Wheeler. (1996). "Polymorphisms in the bovine
185 *beta*-casein 5' flanking region." J. Dairy Sci. **79**: 347-349.

186

187 Bonsing, J., J.M. Ring, A.F. Stewart and A.G. MacKinlay. (1988). "Complete
188 nucleotide sequence of the bovine β -casein gene". Aust J. Biol. Sci. **41**: 527-537.

189

190 Damiani, G., Pilla, F., Leone, P. and Caccio, S. (1992). "Direct sequencing and
191 bidirectional allele specific polymerase chain reaction of the bovine β -casein B
192 variant." Anim. Genet. **23**: 561-566.

193

194 Doppler, W., Welte, T. and Phillipp, S. (1995). "CCAAT/enhancer binding protein
195 isoforms beta and delta are expressed in mammary epithelial cells and bind to
196 multiple sites in the *beta*-casein gene promoter." J. Biol. Chem. **270**: 1854-1862.

197

198 Farrell, Jr, H.M., Jimenez-Flores, R., Bleck, G.T., Brown, E.M., Butler, J.E., Creamer,
199 L.K., Hicks, C.L., Hollar, C.M., Ng-Kwai-Hang, .F., and Swaisgood, H.E. (2004).

200 "Nomenclature of the proteins of cows' milk-sixth edition." J. Dairy Sci. **87**. 1641-
201 1674.

202

203 Inman, C.K., Li, N. and Shore, P. 2005. Oct-1 counteracts autoinhibition of Runx2
204 DNA binding to form a novel Runx2/Oct-1 complex on the promoter of the mammary
205 gland-specific gene β -casein. Mol. Cell. Biol. **25**: 3182-3193.

206

207 Lechner, J., Welte, T., Tomasi, J.K., Bruno, P. and Cairns, C. (1997). "Promoter-
208 dependent synergy between glucocorticoid receptor and Stat5 in the activation of β -
209 casein gene transcription." J. Biol. Chem. **272**: 20954-20960.
210

211 Lien, S., Alestrom, P., Klungland, H. and Rogne, S. (1992). "Detection of multiple
212 β -casein (CASB) alleles by amplification created restriction sites (ACRS)." Anim.
213 Genet. **23**: 333-338.
214

215 Lin, C.Y., McAllister, A.J., Ng-Kwai-Hang, K.F. and Hayes, J.F. (1986). "Effects of
216 milk protein loci on first lactation production in dairy cattle." J. Dairy. Sci. **69**: 704-
217 712.
218

219 Martin, P., M. Szymanowska, L. Zwierzchowski, and C. Leroux. (2002). "The
220 impact of genetic polymorphisms on the protein composition of ruminant milks."
221 Repro. Nutr. Dev. **42**: 433-459.
222

223 Marziali, A.S. and Ng-Kwai-Hang, K.F. (1986). "Relationships between milk
224 protein polymorphisms and cheese yielding capacity." J. Dairy Sci. **69**: 1193-1201.
225

226 Raught, B., Liao, W.S. and Rosen, J.M. (1995). "Developmentally and hormonally
227 regulated CCAAT/enhancer binding protein isoforms influence β -casein gene
228 expression." Mol. Cell Endocrinol. **9**: 1223-1232.
229

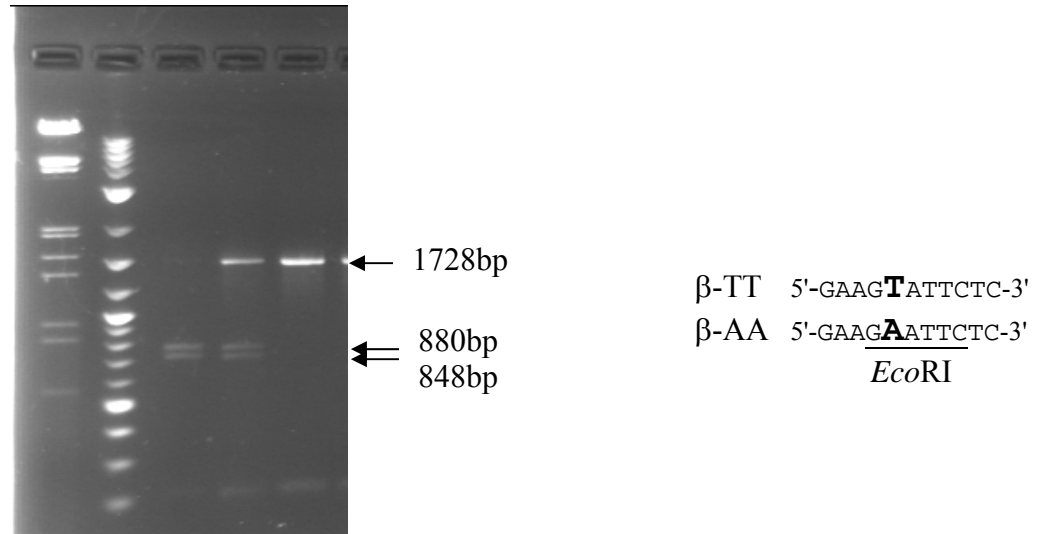
230 Schild, T.A. and Geldermann, H. 1996. Variants within the 5'-flanking regions of
231 bovine milk-protein-encoding genes. III. Genes encoding the Ca-sensitive casein
232 α s1, α s2 and β . Theor. Appl. Genet. **93**: 887-893.
233

234 Stewart, A.F., Bonsing, J., Beattie, C.W., Shah, F., Willis, I.M. and Mackinley, A.G.
235 (1987). "Complete nucleotide sequences of bovine *alpha-s2* and *beta*-casein cDNAs:
236 comparisons with related sequences in other species." Mol. Biol. Evol. **4**: 231-241.
237

238 Szymanowska, M., Siadkowska, E., Lukaszewicz, M. and Zwierzchowski, L. (2004).
239 Association of nucleotide-sequence polymorphism in the 5'-flanking regions of
240 bovine casein genes with casein content in cow's milk. Lait. **84**: 579-590.
241

242 **Figure 1:**

243



244

245 **Table 1.**

246

247

248

249

250

Breed	TT	TA	AA	n
Holstein Friesian	37.5	50	12.5	8
Irish Friesian	75	25	0	4
Dutch Friesian	50	50		4
Norwegian Red	50	50		2
Normande	75	25		4
Montbeliarde	50	25	25	4
Kerry	62.5		37.5	8
Limousin		100		6
Charlois		66	33	2

251 **Table 2.**

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

Animal	β-Casein CDS variant	β-Casein promoter variant
	Dairy Breeds	
	Holstein Friesian	
0011	A1A2	β -TA
0026	A1A2	β -TT
3048	A1A2	β -TA
9615	A2	β -TA
0050	A1A2	β -TT
0059		β -TA
0081	A2	β -AA
0876	A1	β -TT
	Irish Friesian	
0599	A1A2	β -TT
1668	A1A2	β -TT
1270	A2	β -TA
1257	A1A2	β -TT
	Norwegian Red	
0407	A1A2	β -TA
0287	A1	β -TT
	Dutch Friesian	
0188	A1	β -TT
0508	A1A2	β -TA
1535	A1	β -TT
0191	A2	β -TA
	Dual purpose Breeds	
	Normande	
0163	A1	β -TT
0166	A1	β -TT
1226	A1	β -TT
1267	A1A2	β -TA
	Montbeliarde	
1212	A1	β -TT
0130	A2	β -AA
1023	A1A2	β -TA
1545	A1	β -TT
	Kerry	
39	A2	β -AA
40	A2	β -AA
41	A1A2	β -TT
42	A1A2	β -TT
43	A2	β -TT
44	A2	β -AA
45	A1A2	β -TT
46	A2	β -TT
	Beef Breeds	
	Charlois	
0292	A2	β -TA
191C	A1A2	β -TA
	Limousin	
0183	A2	β -AA
0094	A2	β -TA
42L	A1A2	β -TA
215W	A1A2	β -TA
0069	A2	β -AA
0086	A1A2	β -TA

278 Figure legends.

279 **Figure 1. RFLP analysis of the β -casein promoter.**

280 Lane 1: λ HindIII/EcoRI marker. Lane 2: 100bp marker. Lane 3: β -AA 880bp and
281 848bp fragment. Lane 4: β -TA = 1728bp, 880bp + 848bp fragment. Lane 5. β -TT =
282 1728bp fragment.

283

284

285 **Table 1.** The percentage incidence of promoter variants in bovine breeds surveyed.

286 **Table 2.** The incidence of variants in the promoter and coding sequence of the β -
287 casein gene.

288

289

290

291

292