Gene expression phenotypes for lipid metabolism and intramuscular fat in skeletal muscle of cattle


*Australian Cooperative Research Centre for Beef Genetic Technologies (Beef CRC), Armidale, NSW 2351, Australia; †CSIRO Animal, Food and Health Sciences, Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia, QLD 4067, Australia; ‡School of Chemistry and Molecular Biosciences, Faculty of Science, University of Queensland, St. Lucia, QLD 4067, Australia; §CSIRO Food Futures Flagship, Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia, QLD 4067, Australia; and #NSW Department of Primary Industries, Beef Industry Centre, JSF Barker Building, University of New England, Armidale, NSW 2351, Australia

ABSTRACT: Gene expression phenotypes were evaluated for intramuscular fat (IMF) in bovine skeletal muscle as an alternative to traditional estimates of IMF%. Gene expression data from a time course of LM development in high- and low-marbling Bos taurus cattle crosses were compared to identify genes involved in intramuscular adipocyte lipid metabolism with developmentally similar gene expression profiles. Three sets of genes were identified: triacylglyceride (TAG) synthesis and storage, fatty acid (FA) synthesis, and PPARγ-related genes. In an independent analysis in the LM of 48 Bos indicus cattle, TAG and FA gene sets were enriched in the top 100 genes of which expression was most correlated with IMF% (P = 1.2 × 10−24 and 3.5 × 10−9, respectively). In general, genes encoding enzymes involved in the synthesis of FA and TAG in the intramuscular adipocytes were present in the top 100 genes. In B. indicus, effects of a steroid hormone growth promotant (HGP), 2 experimental sites [New South Wales (NSW) and Western Australia (WA)], and 3 tenderness genotypes on the expression levels of genes in the TAG gene set and the correlation of gene expression with IMF% were investigated. Although correlation between expression of 12 individual TAG genes and IMF% was observed in HGP-treated animals in both experimental sites (mean r = 0.43), correlation was not observed for untreated animals at the NSW site (mean r = −0.07, P < 3 × 10−6). However, TAG genes showed an average 1.6-fold (P < 0.0004) reduction in expression in the LM of HGP-treated cattle relative to untreated cattle, an effect consistent across both experimental sites. Cattle possessing the favored tenderness calpain 1 and 3 and calpastatin alleles exhibited a greater (P = 0.008) reduction in expression in NSW (1.8-fold reduction, P = 0.0002) compared with WA (1.2-fold reduction, P = 0.03). Tenderness genotype had no impact (P > 0.05) on the correlation of TAG genes with IMF%. In general, the interactions among genotype, treatment and location, and TAG gene set gene expression were consistent with the interactions among the same factors and IMF% detected using 255 animals, of which the 48 in this study were a subset. Thus, the TAG gene set constitutes a gene expression phenotype able to predict effects of different genotypes and treatments on IMF% using much smaller groups than current approaches, even in animals with very low IMF%.

Key words: beef cattle, fatty acid synthesis, gene expression phenotype, intramuscular fat percentage, tenderness genotype, triacylglyceride synthesis


INTRODUCTION

The juiciness and flavor of beef are 2 important meat quality determinants affected by the amount of intramuscular fat (IMF) present in the muscle
Lipid gene expression in bovine skeletal muscle

(Hocquette et al., 2010). Grain feeding increases IMF% (Williams et al., 1983; Reverter et al., 2003), whereas steroid hormone growth promotant (HGP) treatment reduces IMF% (Bartle et al., 1992; Hunter, 2010). The genotype of cattle also affects IMF%; for example, Bos taurus cattle have an increased marbling potential compared with B. indicus cattle (Wheeler et al., 1994). Intramuscular fat percentage is a complex trait, and generally large numbers of animals are required to demonstrate significant interactions among treatments, genetics, and IMF% (Cafe et al., 2010a; Barendse, 2011; Bolormaa et al., 2011). Genome-wide association studies have had limited success in the identification of the genetic drivers of IMF% (Barendse, 2011; Bolormaa et al., 2011). This may, in part, be due to the different mechanisms through which the same phenotype is manifested. For example, within intramuscular adipocytes alone, among the potential contributing processes to the IMF% phenotype are FA and TAG synthesis and degradation, TAG storage, cell number, and cell volume.

The emergence of comprehensive high-throughput transcriptome analysis methodologies has opened the door to the development of new phenotypes based on well-defined biological processes and characterized by small numbers of genes (Cherel et al., 2012). This approach may provide the means to break complex traits down into a number of simplified component parts. Given the commercial value and complexity of IMF%, it is an appropriate trait to use to evaluate the utility of this approach. In this study, we aimed to identify sets of gene expression markers correlated with IMF% and to explore their use to predict the effect of different environments, treatments, and genotypes on IMF% at the level of groups of animals using small numbers of individuals.

**MATERIALS AND METHODS**

The animal procedures performed in this study were approved by the Industry and Investment New South Wales (NSW) Orange Agricultural Institute Animal Ethics Committee, Commonwealth Scientific and Industrial Research Organization Rockhampton Animal Experimentation Ethics Committee, and the Department of Agriculture and Food, Western Australia (WA) Animal Ethics Committee.

**Animal Resources**

The B. taurus cattle used in this study consisted of a subset from a previously described experiment (Cafe et al., 2010a,b, 2011). The subset included 48 Brahman steers varying in tenderness genotypes, an environment contrast, and an HGP treatment. Of the 48 cattle, there was a balance for HGP treatment (n = 24 for treated and untreated). For experimental site, n = 26 and 22 for NSW and WA, respectively. For tenderness genotype, n = 18, 16, and 14 for tough, intermediate, and tender genotypes, respectively. The tenderness genotypes were based on these 3 genes and SNP: calpastatin (CAST), CAST3–84 (G/A in the 3’ UTR; Barendse, 2002), calpain 3 (CAPN3), CAPN3JK (T/G in an intron; Barendse et al., 2008), and calpain 1 (CAPN1), CAPN1–4751 (T/C in an intron; White et al., 2005). The presence of 2 favorable alleles of each of these genes was shown previously to be associated with an improvement in tenderness (Cafe et al., 2010b). For the purpose of this experiment, a “tough” genotype had no favorable alleles for CAST3–84, CAPN3JK, and CAPN1–4751 (designated genotype G000), and a “tender” genotype had 2 favorable alleles for each of the CAST3–84 and CAPN3JK SNP and at least 1 favorable allele for the CAPN1–4751 SNP (designated genotype G221). The environment contrast was between 2 Australian experimental sites, NSW and WA, described in more detail previously (Cafe et al., 2010a,b). The HGP treatment was a Revalor-H (Virbac, Milperra, NSW, Australia) implant (200 mg trenbolone acetate and 20 mg estradiol) in the ear of the animals in the treated group, with equal numbers of animals treated within each genotype and each site, as previously described in more detail (De Jager et al., 2011).
**IMF% Measures and Analyses**

For the 255 Brahman steers in the full experiment (Cafe et al., 2010a,b), samples of LM were taken immediately after slaughter, and IMF% was determined by near infrared spectrophotometry (Perry et al., 2001). The IMF% values were obtained for 223 animals, including 44 of the 48 animals with samples used in the microarray experiment. Supplementary Table 1 shows the summary statistics for the IMF% across all the animals and by the design effects. The GLM procedure (SAS Inst. Inc., Cary, NC) was used to fit a linear model for the analysis of IMF%. The model included the effects of array (a binary variable indicating whether or not the animal was included in the microarray experiment), HGP or no HGP treatment, experimental site (NSW or WA), and tenderness genotype (tender, intermediate, or tough, as defined earlier for the markers at the CAPN and CAST genes). Solutions from the GLM were then used to adjust the IMF% phenotype of each individual, and the resulting adjusted phenotypes were used in the correlation of IMF% with gene expression, calculated using the CORREL function of Excel (Microsoft, Redmond, WA). For the 4 missing IMF% values, the average of the IMF% across the data set (1.86%) was used (Supplementary Table 1). Correlations were calculated for all genes with IMF% for each of the groups of all animals and for the subsets of only NSW animals, only WA animals, only HGP-treated animals, only control animals, only NSW HGP-treated animals, only WA HGP-treated animals, only NSW control animals, and only WA control animals.

**Total RNA Samples and Microarray**

Gene Expression Data Analysis

The microarray experiment for the development time course of LM in PxH and WxH cattle was described in detail previously (Hudson et al., 2009a,b). Briefly, 2-color hybridizations (1 sample from a PxH and 1 from a WxH individual at the same time point) were undertaken using the bovine Agilent (Santa Clara, CA)

| Table 1. The triacylglyceride (TAG) gene set and additional TAG synthesis and storage genes supported by correlation of gene expression in the LM with intramuscular fat (IMF) percentage in the Brahman cattle or correlation of gene expression across development of the LM in Piedmontese cross Hereford (PxH) and Wagyu cross Hereford (WxH) cattle |
|---------------------------------------------------------------|---------------------------------------------------------------|
| **Gene** | **Description** | **All animals** | **HGP treatment** | **Source** | **Gene set** |
|---------------------------------------------------------------|---------------------------------------------------------------|
| **CIDEC** | cell-death-inducing DFFA-like effector c | 4 | 0.33 | 0.54 | 0.47 | −0.40 | Network analysis | TAG |
| **DGAT2** | diacylglycerol O-acyltransferase homolog 2 | 6 | 0.32 | 0.47 | 0.42 | −0.03 | Network analysis | TAG |
| **THRSP** | thyroid hormone responsive | 13 | 0.30 | 0.41 | 0.45 | −0.03 | Module 2 Always Corr | TAG |
| **PLIN1** | perilipin 1 | 15 | 0.30 | 0.49 | 0.43 | −0.07 | Module 1 Always Corr | TAG |
| **PKC1** | phospho(enol)pyruvate carboxykinase 1 | 20 | 0.29 | 0.35 | 0.35 | 0.22 | Network analysis | TAG |
| **AGPAT2** | 1-acylglycerol-3-phosphate O-acyltransferase 2 | 21 | 0.28 | 0.42 | 0.06 | 0.10 | Network analysis | TAG |
| **ACSM1** | acyl-CoA synthetase medium-chain family member 1 | 35 | 0.26 | 0.41 | 0.61 | −0.04 | Network analysis | TAG |
| **CIDEC** | cell-death-inducing DFFA-like effector c | 41 | 0.26 | 0.47 | 0.42 | −0.06 | Module 1 Always Corr | TAG |
| **ADIPQ** | adiponectin | 46 | 0.25 | 0.33 | 0.37 | −0.10 | Module 1 Always Corr | TAG |
| **FABP4** | fatty acid binding protein 4, adipocyte | 54 | 0.24 | 0.46 | 0.42 | −0.10 | Module 1 Always Corr | TAG |
| **S100G** | S100 calcium binding protein G | 61 | 0.24 | 0.47 | 0.18 | −0.26 | Module 2 Always Corr | TAG |
| **TUSC5** | tumor suppressor candidate 5 | 73 | 0.23 | 0.38 | 0.46 | −0.07 | Module 1 Always Corr | TAG |
| **ADIG** | adipogenin | 183 | 0.18 | 0.25 | 0.29 | 0.02 | Module 2 Always Corr | TAG |
| **PLSI** | plasmin 1 | 502 | 0.12 | 0.35 | −0.1 | −0.01 | Module 1 Always Corr | None |
| **GPAM** | glycerol-3-phosphate acyltransferase | 1364 | 0.07 | 0.23 | −0.12 | 0.12 | Correlation with IMF% | None |

1Full list of genes and probes is included as Supplementary Table 3.
2All 48 animals were used to calculate the correlation between gene expression and IMF%.
3Hormone growth promotant (HGP); Yes indicates implanted with HGP, and No indicates not implanted.
4Gene rank of the 19,265 probes tested for correlation of expression with IMF% across all 48 Brahman cattle.
5The 12 animals from NSW that were not treated with HGP were used to calculate the correlation of gene expression with IMF%.
6The 24 HGP-treated animals from Western Australia (WA) and New South Wales (NSW) were used to calculate the correlation of gene expression with IMF%.
7The 12 animals from WA that were not treated with HGP were used to calculate the correlation of gene expression with IMF%.
8Genes with the source “Module 1 Always Corr” were members of the first of the lipid metabolism modules in the always correlated landscape (Hudson et al., 2009b). Genes with the source “Module 2 Always Corr” were members of the second of the lipid metabolism modules in the always correlated landscape (Hudson et al., 2009b). Genes with the source “Network analysis” were identified as additional genes having correlated gene expression patterns across with the genes in modules 1 and 2 of the always correlated landscape during development of the LM of PxH and WxH cattle (Hudson et al., 2009b). Genes with the source “Correlation with IMF%” were not identified in the network analysis undertaken in this work and were not present in either of the always correlated landscape lipid metabolism modules (Hudson et al., 2009b).
9DFFA = DNA fragmentation factor
microarray platform (G2519F V1: 015354). Three RNA samples were used for each time point for the prenatal and birth LM samples, and 4 RNA samples were used for each time point for the other postnatal LM samples. Microarray data were normalized using a linear mixed ANOVA model as previously described (Hudson et al., 2009a). The data and networks derived from the data are available as part of the supplementary information linked to the publications (Hudson et al., 2009a,b). The microarray experiment for the Brahman cattle was described in detail previously (De Jager et al., 2011). Briefly, single-color hybridizations were undertaken using the bovine Agilent microarray platform (G2519F V1: 015354). Microarray data were normalized using a linear mixed ANOVA model as previously described (De Jager et al., 2011). The data are available in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus database with the accession number GSE25005. Genes were assigned to probes on the bovine Agilent microarray platform by alignment of the probes to the Btau4.1 and UMD3.1 versions of the bovine genome assembly and to the NCBI bovine and human RefSeq collections using a consensus method as described previously (Mariasegaram et al., 2010).

**Identification of Lipid Metabolism Gene Sets**

Starting with 2 modules of genes involved in lipid metabolism identified previously in the bovine LM "always correlated" landscape (Hudson et al., 2009b; Table 1), a set of gene selection parameters combining differences in the level of gene expression, both within crosses between time points and between crosses at the same time points, was constructed. The following abbreviations are used below: for the crosses, P (PxH) and W (WxH), and for the sampling time points, 60, 135, 195, and 280 (for days post conception), and 3, 7, 12, 20, 25, and 30 (for months after birth). All gene expression values described below are expressed as log base 2. The gene selection parameters are expressed as 1 cross and time point minus the second cross and time point and the numerical selection parameter. They are shown also diagrammatically in Fig. 1. All differences were calculated using the mean values of the expression of the genes from the 3 (prenatal and birth) or 4 (postnatal) samples for each cross. To reduce the frequency of false-negatives, neither the variance of the gene expression values at a single time point nor the significance of the differences between means were taken into account in generating the gene selection parameters.

The gene selection parameter set for the triacylglyceride (TAG) synthesis and storage gene set were determined empirically to include the genes in both modules 1 and 2 from the always correlated landscape (except for *plastin1*), but without including a large number of other genes (Table 1). For the TAG synthesis and storage gene set, the gene selection parameter set was as follows: (W135 minus W60 > 1.5) and (P135 minus P60 > 3.5) and (P135 minus W135 > 0.5) and (P195 minus P135 < 1) and [(P12 minus P7 < –0.5) or (W12 minus W7 < –0.5)] and (P30 minus W30 < –0.5). Two further sets of genes were identified as described below.

For the fatty acid (FA) synthesis gene set, the gene selection parameters were selected empirically to exclude all the genes in the TAG gene set but to include genes with gene expression profiles across development similar to the genes in the TAG gene set (but with a reduced dynamic range in expression changes between time points and differences between crosses), again without generating a very large gene set. The gene selection parameters were as follows: (P135 minus P60 > 0.5 and ≤ 3.5) and (P135 minus W135 > 0.5) and (P195 minus P135 < 0.5) and [(P280 minus P195 < 1)] and (W12 minus W7 < –0.5) or (P12 minus P7 < –0.5) and (P30 minus W30 < –0.5).

For the PPARγ (PPARG) gene set, the gene selection parameters were selected empirically to exclude all the genes in the TAG and FA gene sets but to include genes with gene expression profiles across development

![Figure 1. Diagrammatic representation of the gene selection parameters used to define the gene sets in the gene expression data across the development of the LM of Piedmontese cross Hereford (PxH) and Wagyu cross Hereford (WxH) cattle. Solid lines represent PxH cattle, and dashed lines represent WxH cattle comparisons between time points. Vertical dotted lines represent comparisons between the 2 crosses at the same time point. All values are log base 2 differences in expression, and where calculated, between time points are the later time point value minus the earlier time point value. The peak of primary and secondary myogenesis, the start of adipogenesis (Du et al., 2010), birth, weaning, and entry into feedlot are indicated. TAG = triacylglyceride gene set; FA = fatty acid gene set; PPARG = PPARγ gene set.](image-url)
similar to the genes in the TAG and FA gene sets (but with a reduced dynamic range in expression changes between time points and differences between crosses and no emphasis on an effect at weaning, the 7-mo sample), again without generating a very large gene set. The gene set selection parameters were as follows: (W135 minus W60 > –1) and (P135 minus P60 > 0.5) and (P135 minus W 135 > 0.5) and (W195 minus W135 < 0.5) and (P280 minus P195 < 1) and (P12 minus P7 ≥ –0.5) and (W12 minus W7 ≥ –0.5) and (P25 minus W25 < –0.5) and (P30 minus W30 < –0.5). The full list of genes in the TAG, FA, and PPARG gene sets and the microarray probes used are listed in Supplementary Table 2.

**Bioinformatics and Statistics**

Gene enrichment analysis was undertaken using GOrilla (Eden et al., 2007, 2009). The P-values and the false discovery rate (FDR) Q values were calculated using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995) and were provided in the results output of the GOrilla Web site (http://cbl-gorilla.cs.technion.ac.il/).

The significance of the differences in correlations of gene expression with IMF% of the genes in the TAG gene set in groups of animals was calculated using the paired 2-tailed Student’s t test function in Microsoft Excel. Using this approach, for each of the 6 comparisons explored [HGP vs. control, NSW vs.

Table 2. The fatty acid synthesis (FA) and PPARγ gene sets and additional FA genes supported by correlation of gene expression in the LM with intramuscular fat (IMF) percentage in the Brahman cattle and/or correlation of gene expression across development of the LM in Piedmontese cross Hereford and Wagyu cross Hereford cattle

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Description</th>
<th>Rank4</th>
<th>r</th>
<th>Yes WA + NSW5</th>
<th>No WA6</th>
<th>NSW7</th>
<th>Gene set</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCD</td>
<td>stearoyl-CoA desaturase (delta-9-desaturase)</td>
<td>9</td>
<td>0.30</td>
<td>0.49</td>
<td>0.43</td>
<td>–0.31</td>
<td>FA</td>
</tr>
<tr>
<td>MAL2</td>
<td>mal, T-cell differentiation protein 2</td>
<td>33</td>
<td>0.26</td>
<td>0.50</td>
<td>0.37</td>
<td>0.05</td>
<td>FA</td>
</tr>
<tr>
<td>ELOVL6</td>
<td>ELOVL fatty acid elongase 6</td>
<td>34</td>
<td>0.26</td>
<td>0.41</td>
<td>0.31</td>
<td>–0.02</td>
<td>FA</td>
</tr>
<tr>
<td>FASN</td>
<td>fatty acid synthase</td>
<td>39</td>
<td>0.26</td>
<td>0.40</td>
<td>0.46</td>
<td>–0.30</td>
<td>FA</td>
</tr>
<tr>
<td>ACACA</td>
<td>acetyl-CoA carboxylase alpha</td>
<td>56</td>
<td>0.24</td>
<td>0.41</td>
<td>0.40</td>
<td>–0.21</td>
<td>FA</td>
</tr>
<tr>
<td>CLU</td>
<td>clusterin</td>
<td>65</td>
<td>0.23</td>
<td>0.40</td>
<td>0.18</td>
<td>–0.16</td>
<td>FA</td>
</tr>
<tr>
<td>PCK2</td>
<td>phosphoenolpyruvate carboxykinase 2 (mitochondrial)</td>
<td>92</td>
<td>0.22</td>
<td>0.41</td>
<td>0.46</td>
<td>–0.12</td>
<td>FA</td>
</tr>
<tr>
<td>RBP4</td>
<td>retinol binding protein 4, plasma</td>
<td>139</td>
<td>0.19</td>
<td>0.41</td>
<td>0.16</td>
<td>–0.16</td>
<td>FA</td>
</tr>
<tr>
<td>PDE3B</td>
<td>phosphodiesterase 3B, cGMP-inhibited</td>
<td>206</td>
<td>0.17</td>
<td>0.31</td>
<td>0.09</td>
<td>0.09</td>
<td>FA</td>
</tr>
<tr>
<td>BHMT2</td>
<td>betaine–homocysteine S-methyltransferase 2</td>
<td>207</td>
<td>0.17</td>
<td>0.33</td>
<td>0.50</td>
<td>–0.34</td>
<td>FA</td>
</tr>
<tr>
<td>INSIG1</td>
<td>insulin induced gene 1</td>
<td>412</td>
<td>0.14</td>
<td>0.21</td>
<td>0.31</td>
<td>–0.17</td>
<td>FA</td>
</tr>
<tr>
<td>FBP1</td>
<td>fructose-1,6-bisphosphatase 1</td>
<td>706</td>
<td>0.10</td>
<td>0.36</td>
<td>0.09</td>
<td>–0.13</td>
<td>FA</td>
</tr>
<tr>
<td>ANPEP</td>
<td>alanyl (membrane) aminopeptidase</td>
<td>2091</td>
<td>0.04</td>
<td>0.26</td>
<td>–0.16</td>
<td>–0.15</td>
<td>FA</td>
</tr>
<tr>
<td>ACSS2</td>
<td>acyl-CoA synthetase short-chain family member 2</td>
<td>19</td>
<td>0.29</td>
<td>0.40</td>
<td>0.48</td>
<td>0.01</td>
<td>PPARG</td>
</tr>
<tr>
<td>CYB5A</td>
<td>cytochrome b5 type A (microsomal)</td>
<td>64</td>
<td>0.23</td>
<td>0.31</td>
<td>0.45</td>
<td>–0.10</td>
<td>PPARG</td>
</tr>
<tr>
<td>PPARG</td>
<td>peroxisome proliferator-activated receptor gamma</td>
<td>117</td>
<td>0.20</td>
<td>0.25</td>
<td>0.38</td>
<td>0.18</td>
<td>PPARG</td>
</tr>
<tr>
<td>ACSS3</td>
<td>acyl-CoA synthetase short-chain family member 3</td>
<td>137</td>
<td>0.19</td>
<td>0.36</td>
<td>0.26</td>
<td>0.11</td>
<td>PPARG</td>
</tr>
<tr>
<td>ACLY</td>
<td>ATP citrate lyase</td>
<td>148</td>
<td>0.19</td>
<td>0.31</td>
<td>0.25</td>
<td>–0.19</td>
<td>PPARG</td>
</tr>
<tr>
<td>TKT</td>
<td>transketolase</td>
<td>229</td>
<td>0.17</td>
<td>0.29</td>
<td>0.29</td>
<td>–0.23</td>
<td>PPARG</td>
</tr>
<tr>
<td>TF</td>
<td>transferrin</td>
<td>310</td>
<td>0.15</td>
<td>0.16</td>
<td>0.58</td>
<td>0.03</td>
<td>PPARG</td>
</tr>
<tr>
<td>IDH1</td>
<td>isocitrate dehydrogenase 1 (NADP+), soluble</td>
<td>320</td>
<td>0.15</td>
<td>0.21</td>
<td>0.36</td>
<td>–0.27</td>
<td>PPARG</td>
</tr>
<tr>
<td>QPRT</td>
<td>quinolinate phosphoribosyltransferase</td>
<td>329</td>
<td>0.15</td>
<td>0.14</td>
<td>0.36</td>
<td>–0.15</td>
<td>PPARG</td>
</tr>
<tr>
<td>ALAD</td>
<td>aminolevulinic dehydratase</td>
<td>422</td>
<td>0.13</td>
<td>0.23</td>
<td>–0.26</td>
<td>0.13</td>
<td>PPARG</td>
</tr>
<tr>
<td>CEBPA</td>
<td>CCAAT/enhancer binding protein (C/EBP), α</td>
<td>593</td>
<td>0.11</td>
<td>0.16</td>
<td>0.09</td>
<td>0.003</td>
<td>PPARG</td>
</tr>
<tr>
<td>ACER3</td>
<td>alkaline ceramidase 3</td>
<td>683</td>
<td>0.10</td>
<td>0.12</td>
<td>0.14</td>
<td>–0.08</td>
<td>PPARG</td>
</tr>
<tr>
<td>G6PD</td>
<td>glucose-6-phosphate dehydrogenase</td>
<td>982</td>
<td>0.08</td>
<td>0.17</td>
<td>0.25</td>
<td>–0.19</td>
<td>PPARG</td>
</tr>
<tr>
<td>GSTA1</td>
<td>glutathione S-transferase alpha 1</td>
<td>998</td>
<td>0.08</td>
<td>0.20</td>
<td>–0.07</td>
<td>–0.14</td>
<td>PPARG</td>
</tr>
<tr>
<td>INTS9</td>
<td>integrator complex subunit 9</td>
<td>1183</td>
<td>0.07</td>
<td>0.18</td>
<td>0.08</td>
<td>–0.10</td>
<td>PPARG</td>
</tr>
<tr>
<td>HSD17B12</td>
<td>hydroxysteroid (17-β) dehydrogenase 12</td>
<td>74</td>
<td>0.23</td>
<td>0.41</td>
<td>0.06</td>
<td>–0.13</td>
<td>None</td>
</tr>
<tr>
<td>PTPLB</td>
<td>patatin-like phospholipase domain containing 4</td>
<td>275</td>
<td>0.16</td>
<td>0.35</td>
<td>0.05</td>
<td>–0.12</td>
<td>None</td>
</tr>
</tbody>
</table>

1Full list of genes and probes is included as Supplementary Table 3.
2All 48 animals were used to calculate the correlation between gene expression and IMF%.
3Hormone growth promotant (HGP); Yes indicates implanted with HGP, and No indicates not implanted.
4Gene rank of the 19,265 probes tested for correlation of expression with IMF% across all 48 animals.
5The 24 HGP treated animals from Western Australia (WA) and New South Wales (NSW) were used to calculate the correlation of gene expression with IMF%.
6The 12 animals from WA that were not treated with HGP were used to calculate the correlation of gene expression with IMF%.
7The 12 animals from NSW that were not treated with HGP were used to calculate the correlation of gene expression with IMF%.
WA, HGP vs. control (WA), HGP vs. control (NSW), G221 (HGP) vs. G220 (HGP), G221 (HGP) vs. G000 (HGP), G220 (HGP) vs. G000 (HGP)], we calculated the correlation of expression of each of the 12 genes in the TAG gene set (CIDEA, ADIG, S100G, PCK1, PLIN1, FABP4, ADIPOQ, PLS1, AGPAT2, DGAT2, CIDEC, and TUSC5; full gene names are listed in Table 1) with IMF% and paired genes between the 2 groups of animals.

The significance of the differences in expression levels of the genes in the TAG gene set between pairs of conditions (such as HGP-treated and control animals) were calculated using the unpaired 2-tailed Student’s *t* test function in Microsoft Excel. To avoid possible violations of the distributional assumptions implied in the *t* test, we also undertook a nonparametric approach to ascertain the statistical significance of the observed expression patterns. Using this approach, for each of the 6 comparisons explored [HGP minus control (WA) vs. HGP minus control (NSW), HGP minus control (tender) vs. HGP minus control (tough), tender minus tough (WA) vs. tender minus tough (NSW), tender minus tough (HGP) vs. tender minus tough (control), NSW minus WA (HGP) vs. NSW minus WA (control), and NSW minus WA (tender) vs. NSW minus WA (tough)], we measured the average differential expression of 12 genes in the TAG gene set (CIDEA, ADIG, S100G, PCK1, PLIN1, FABP4, ADIPOQ, PLS1, AGPAT2, DGAT2, CIDEC, and TUSC5; full gene names are listed in Table 1) and compared it with that obtained from a random sample of 12 genes. This process was repeated 1 million times, and a nonparametric estimate of the significance was obtained from the number of occasions in which the average differential expression of 12 random genes exceeded that of the 12 genes in the TAG gene set included in the analysis.

Genes supported by the integrated analysis were defined as those with values for the correlation of gene expression in LM with IMF% and their differential gene expression, expression in WxH LM minus expression in PxH LM, larger than the values defined by the relationship (differential gene expression in LM WxH minus PxH at 25 mo) = 0.07 × (correlation of gene expression in LM in HGP treated Brahmans with IMF%)$^{-1.25}$ or with values less than those defined by the equivalent equation for the reciprocal relationships. The parameters for the equation were estimated empirically on the basis of the distribution of genes in the TAG and FA gene sets between the 2 sides.

## RESULTS

### Identification of TAG Synthesis and Storage Gene Set

To identify sets of genes potentially suitable for use as a gene expression phenotype(s) for lipid metabolism, we started with 2 small modules (groups of genes with correlated expression data sets and with gene expression profiles distinct from all other groups of genes) of genes involved in lipid metabolism identified in the bovine LM development always correlated landscape (Hudson et al., 2009b). These 2 modules of genes were designated “Module 1 Always Corr” (6 genes) and “Module 2 Always Corr” (3 genes) in the source column (Table 1). In brief, the always correlated landscape was built using data from a time course of gene expression from 60 d after conception to 30 mo of age (Fig. 1) in a high-IMF% cross (WxH) and a high-muscling, low-IMF% cross (PxH) and a smaller number of Brahman animals. The time course data set included samples taken during the peaks of primary and secondary myogenesis, at birth, before and after weaning, during grass feeding, and after a period in the feedlot (Fig. 1). Critical time points for adipogenic processes in the muscle in this experiment were initiation of adipogenesis around d 120 of prenatal development (Du et al., 2010), weaning, increasing age, and entry into feedlot. Although identified as separate modules in the original analysis (Hudson et al., 2009b), the expression patterns of the 2 sets of genes were very similar across development from 60 d after conception to 30 mo of age (Fig. 2). We then elected to expand the modules by constructing a set of gene selection parameters focusing on the key time points in intramuscular adipogenesis. The selection parameters were designed to capture all of the genes present in the original modules, as well as some additional genes, without a very large increase in the number of genes in the module. The selection parameters used the differences in gene expression level between the 2 crosses at the same time points (Fig. 1). In addition, the selection parameters also used the differences in gene expression level between adjacent time points independently within each of the 2 crosses (Fig. 1). Five additional genes were identified as having a similar expression profile (Fig. 2A) to the genes in the original modules, source “Network analysis” in Table 1. Plastin 1 was not included in the set of genes identified using the selection parameters, as its inclusion generated a very large set of additional genes (data not shown). As a large number of the products of the genes in the enlarged gene set are involved in the synthesis or storage of TAG (Fig. 3), the gene set was named the “TAG synthesis and storage gene set,” abbreviated as the “TAG gene set.”
Figure 2. Gene expression profiles of selected genes across the development of the LM of Piedmontese cross Hereford (PxH) and Wagyu cross Hereford (WxH) cattle. (A) Representative gene from the triacylglyceride (TAG) synthesis and storage gene set, thyroid hormone responsive (THRSP). (B) Representative gene from the fatty acid (FA) synthesis gene set, acetyl-CoA carboxylase α (ACACA). (C) Representative gene from the PPARγ (PPARG) gene set, PPARG. (D) Glycerol-3-phosphate acyltransferase (GPAM). (E) Hydroxysteroid (17-β) dehydrogenase 12 (HSD17B12). (F) Patatin-like phospholipase domain containing 4 (PTPLB). In each plot data from the WxH animals are shown with a dashed line, and PxH animals are shown with a solid line.
Figure 3. The proposed fatty acid (FA) and triacylglyceride (TAG) synthesis and storage pathways in skeletal muscle adipocytes of cattle. Genes with black backgrounds are in the TAG gene set. Genes with dark gray backgrounds are in the FA gene set. Genes with light gray backgrounds are in the PPARγ gene set. Genes with white backgrounds are not in any of the gene sets but are supported by the analysis undertaken here. Question marks indicate steps for which there is little support in the current data sets for any of the known genes encoding proteins catalyzing these steps, but the most likely gene based on the available evidence is indicated.
Identification of FA synthesis and PPARG Gene Sets

Two further sets of genes involved in lipid biosynthetic processes and with similar but differentiable gene expression patterns in the development data set to the TAG gene set genes also were identified using the same approach as selecting subsets of genes from the full data set (Fig. 1). However, these gene sets were not initiated using existing modules from the always correlated landscape (Hudson et al., 2009b). These gene sets were designated the FA synthesis (Figs. 2B and 3) and the PPARG gene sets (Fig. 2C). The FA (25 genes) and PPARG (43 genes) gene sets were differentiated from the TAG gene set primarily on the much smaller increase in expression at d 135 in the WxH animals compared with the PxH animals, and the PPARG gene set was differentiated from both the TAG and FA gene sets on the lack of a discernible peak in expression in the samples from the LM of 7-mo-old animals collected at weaning.

Correlation of Gene Expression with IMF% in Brahman Animals

Gene expression data from microarrays are noisy, potentially leading to high false-positive and false-negative rates. Indeed, although the TAG gene set contained many genes with known roles in lipid metabolism, this was not the case for the genes in the PPARG gene set. To validate and increase the robustness of the gene sets, comparison with the data from a second independent experiment analyzed with a different method was undertaken.

To generate the independent comparison data set, gene expression data and IMF% values for the LM were obtained from 48 Brahman individuals. The correlation between the expression levels of all of the probes on the microarray and the IMF% was calculated (Supplementary Table 3). Gene ontology (GO) enrichment analysis for the top 100 genes with expression most positively correlated with IMF% identified the term “triglyceride metabolic process” as the most enriched \( P < 4 \times 10^{-9}, \) FDR \( Q < 4 \times 10^{-5} \), followed closely by many other terms related to lipid metabolic processes. In contrast, no significant enrichment in any GO term was observed in the 100 genes with expression most negatively correlated with IMF% \( P > 0.0006, \) FDR \( Q = 1 \). The genes in the TAG and FA but not PPARG gene sets also were enriched greatly toward the top of the list (Tables 1 and 2) and particularly in the top 100 \( P = 1.2 \times 10^{-24} \) and \( P = 3.5 \times 10^{-9} \), respectively, for the TAG and FA gene sets). When the correlations between gene expression levels and IMF% were calculated separately for the HGP-treated and the untreated animals, a large difference in correlation was observed for the genes in the TAG (average correlation of 12 TAG gene set genes in HGP-treated animals = 0.41, and in untreated animals \( r = 0.10, P = 7.8 \times 10^{-6} \); Fig. 4). No similar effects were observed when the data were split on the basis of experimental site \( P = 0.13 \); however, when the animals were divided by both HGP treatment and experimental site, enrichment of TAG genes in the top 100 most correlated genes was no longer observed in the untreated animals from the NSW experimental site (Tables 1 and 2). An average correlation of 0.43 between the expression of 12 TAG gene set genes and IMF% was observed in HGP-treated animals combining both experimental sites. In contrast, a mean correlation of –0.07 was observed for untreated animals at the NSW site \( P = 2.8 \times 10^{-6} \). When HGP-treated animals were grouped on the basis of their tenderness genotype (combining animals from the 2 experimental sites to avoid very small group sizes), TAG gene set genes were correlated with IMF% in all 3 groups. Average correlations for the 12 genes in the TAG gene set with IMF% were 0.5 (G221), 0.38 (G220), and 0.45 (G000); for all 3 pairwise comparisons \( P > 0.05 \). Similar relationships were observed with the FA gene set (data not shown).

Integrated Data Analysis

The IMF% in the 48 Brahmans was very low to low (0.57% to 3.4%, average 1.9%), and correlation coefficients were low, potentially generating a large number of false-positive correlations. In contrast, in the WxH and PxH individuals, the IMF% values were greater, and differences between the 2 crosses were large; mean IMF% values were 8.82% for WxH and 5.05% for PxH \( P < 0.05 \); Greenwood et al., 2006), but only 4 animals of each cross were analyzed in the gene expression experiment (Hudson et al., 2009b), rendering correlations unreliable. In addition, given the differences between the WxH and PxH animals, gene expression differences between the 2 crosses may have been due to other factors in addition to the differences in IMF% (e.g., differences in the activity of the MSTN gene affecting other traits; Greenwood et al., 2006). To overcome the probable limitations of each of the 2 data sets, correlation of gene expression with IMF% in HGP-treated animals was plotted against difference in expression between WxH and PxH animals at 25 mo (Fig. 4B). Only the HGP-treated animals were used in this analysis, as correlation between genes in the TAG, FA, and PPARG gene sets in the NSW animals not treated with HGP was not enriched, and inclusion of this data set may have weakened the discriminating power of the correlation data set. The 25-mo time point was chosen to include a grass-fed vs. feedlot comparison, increasing the selection power of the analysis. The genes in the intersection are predicted to be highly likely to be involved in IMF deposition, independent of diet and breed or subspecies. Genes
encoding proteins catalyzing most of the steps involved in the synthesis of long-chain saturated and unsaturated FA, up to C18, via the FASN/endoplasmic reticulum pathway and the synthesis and storage of TAG were present in the region of significance (Fig. 3). The expression of glycerol-3-phosphate acyltransferase (GPAM; TAG synthesis) and hydroxysteroid (17-β) dehydrogenase 12 (HSD17B12) and patatin-like phospholipase domain containing 4 (PTPLB; FA synthesis) also lay outside the bulk of the distribution (Fig. 4B), supporting their roles in the FA and TAG pathways in intramuscular adipocytes. However, the expression profile of GPAM was quite distinct from the representative TAG gene set expression profile (Fig. 2D), albeit possessing features observed in the representative FA and PPARγ gene set expression profiles. The HSD17B12 expression profile was also distinct from the representative FA gene set expression profile (Fig. 2E), whereas the PTPLB expression profile was similar to the representative PPARγ gene set expression profile (Fig. 2F), but not enough to have been included in the PPARγ gene set. In contrast, none of the genes encoding proteins with 1,2-diacyl-sn-glycerol synthesis activity (lipins 1–3, LPIN1, LPIN2, LPIN3) and phosphatidic acid phosphatases (PPAP) type 2A-C (PPAP2A, PPAP2B, and PPAP2C) lay outside the bulk of the distribution. The genes encoding proteins with trans-2,3-enoyl-CoA reductase activity [the endoplasmic reticulum trans-2,3-enoyl-CoA reductase (TECR), the mitochondrial trans-2-enoyl-CoA reductase (MECR), the peroxisomal trans-2-enoyl-CoA reductase (PECR), and trans-2,3-enoyl-CoA reductase-like (TECRL or SRD5A2L)] lay outside the bulk of the distribution. Unexpectedly, this was except for the gene encoding the periplasmic enzyme (PECR) when correlation with HGP-treated animals only was used. In addition, none of the genes encoding other proteins involved in the elongation of long-chain and very long-chain fatty acids (ELOVL fatty acid elongases 1, 2, 4, and 7; neither ELOVL fatty acid elongase 3 nor ELOVL fatty acid elongase 5 were tested) or their desaturation [stearoyl-CoA desaturase 5 (SCD5), fatty acid desaturases 1–3, 6 (FADS1–3,6; Guillou et al., 2010)] were located outside of the bulk of the distribution. None of the genes encoding proteins involved in the elongation of FA in the mitochondria, acetyl-CoA acyltransferase 2 (ACAA2), hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), α (HADHA) and β subunits (HADHB), mitochondrial enoyl CoA hydratase short chain 1 (ECHS1), and MECR, were located outside of the bulk of the distribution. Space limitations prevent a discussion of the genes outside the bulk of the distribution and not present in the TAG, FA, and PPARγ gene sets. Results of this analysis will be published separately.

Figure 4. The distribution of correlations between the expression of genes in the LM of Brahman cattle and intramuscular fat (IMF) percentage and differential gene expression at 25 mo of age in the LM of Piedmontese cross Hereford (PxH) and Wagyu cross Hereford (WxH) cattle. (A) Comparison between hormone growth promotant (HGP) treated and control Brahman animals. Triacylglyceride (TAG) gene set genes are shown by circles. (B) Plot of correlation of gene expression in HGP-treated Brahman cattle with IMF% against differential expression at 25 mo between WxH and PxH cattle. TAG gene set genes (circles), fatty acid (FA) gene set genes (triangles), and PPARγ (PPARγ) gene set genes (pluses) are shown. The locations of glycerol-3-phosphate acyltransferase (GPAM), hydroxysteroid (17-β) dehydrogenase 12 (HSD17B12), and patatin-like phospholipase domain containing 4 (PTPLB) are indicated (squares). In the quadrant containing the data points for genes with expression positively correlated with IMF% and more highly expressed in WxH than PxH the line for \( y = 0.07x^{-1.25} \) and the equivalent line in the negative quadrant are shown. These lines were used to discriminate between genes supported and genes not supported by the integrated analysis. For ease of display 3 points with differential expression between –2.5 and –4.5 have been omitted from the diagram.
Evaluating Published Correlations of Gene Expression with IMF% or Marbling

The expression of many genes or their products in muscle has been reported to be positively correlated with IMF% or marbling in cattle: ADIPOQ, SCD, and THRSP (Wang et al., 2009b), G6PD and leptin (LEP; Graugnard et al., 2009), FABP4 (Lee et al., 2008; Pickworth et al., 2011), ADAM metallopeptidase with thrombospondin type 1 motif, 4 (ADAMTS4), and transforming growth factor, β1 (TGFβ1; Lee et al., 2010), coiled-coil domain containing 3 (CCDC3; Eberlein et al., 2010), IL-1 receptor-associated kinase 1 (IRA1; Eberlein et al., 2011), CEBPA and PPARg (Kim et al., 2011; Lim et al., 2011), delta-like 1 homolog (DLK1; Pickworth et al., 2011), runt-related transcription factor 1; translocated to chromosome 1 (cyclin D-related; RUNX1T1; Lim et al., 2011), GPAM, ACACA, FASN, LPL, CD36 molecule (thrombospondin receptor), solute carrier family 27 (fatty acid transporter) member 27 (fatty acid transporter) member 1 [SLC27A1, also known as (aka) FATP1], 1-acylglycerol-3-phosphate O-acyltransferase 1 (lysocephatidic acid acyltransferase, α; AGPAT1), diacylglycerol O-acyltransferase 1 (DGAT1), and DGAT2 (Jeong et al., 2012). An additional set of genes has been shown to be negatively correlated with IMF% or marbling, including titin (TTN), nebulin (NEB; Lee et al., 2008), calcium/calmodulin-dependent protein kinase II α (CAMK2A; Lee et al., 2008; Lim et al., 2011), retinoid X receptor α (RXRA; Lim et al., 2011), heat shock 27 kDa protein 1 (HSPB1), TNF receptor superfamily member 6 (FAS), angiotensinogen [serpin peptidase inhibitor clade A member 8 (AGT); Kim et al., 2011], patatin-like phospholipase domain containing 2 (PNPLA2, aka ATGL), monoglyceride lipase (MGLL, aka MGL), acyl-CoA dehydrogenase very long chain (ACADVL, aka VLCAD), and acyl-CoA dehydrogenase C-4 to C-12 straight chain (ACADM, aka MCD; Jeong et al., 2012). In our analysis of cattle muscle gene expression, only the positively correlated genes ACACA, ADIPOQ, CCDC3, CEBPA, DGAT2, G6PD, GPAM, FABP4, FASN, LPL, PPARg, SCD, and THRSP from the list above were well supported [i.e., genes were above the line defined by the relationship (differential gene expression WxH minus PxH at 25 mo) = 0.07 × (correlation of gene expression with IMF%)^{1.25} or below the equivalent line in the reciprocal quadrant (Fig. 4B)]. None of the unsupported genes were included in the TAG, FA, or PPARg gene sets.

The expression of pinin (PNN)-interacting serine/arginine-rich protein (PNNISR aka SFRS18; Wang et al., 2009a), eukaryotic translation elongation factor 1 α 2 (EEF1A2), fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor; FABP3), low-density lipoprotein receptor (LDLR), matrix Gla protein (MGP), obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF (OBSCN), pyruvate dehydrogenase (lipoamide) δ (PDHB), triadin (TRDN), and ryanodine receptor 1 (skeletal; RYR1; Serao et al., 2011) in the LM has been reported to be positively correlated with IMF% in pigs, but none of these genes was supported by our data from cattle muscle. The unsupported genes were also not included in the TAG, FA, or PPARg gene sets.

Using the TAG Gene Set to Predict the Effect of HGP Treatment, Site, and Calpain and Calpastatin Genotype on IMF%

Twelve of the 14 genes in the TAG gene set were in the top 100 genes with expression correlated with IMF%. This was a much greater proportion than for the FA and the PPARg gene sets. Thus, the genes in the TAG gene set appeared to be the best source of genes to use in a gene expression phenotype for IMF%. Using multiple genes was expected to reduce the influence of noise and individual variation and therefore increase the robustness and significance of the signal. The utility of a subset of the TAG gene set for the prediction of the effects of different treatments or genetics on IMF% at the level of groups of animals was investigated. In all of the comparisons analyzed (Figs. 5A–5F), the average differential expression of the 12 genes used in the TAG gene set (CIDEA, ADIG, S100G, PCK1, PLIN1, FABP4, ADIPOQ, PLS1, AGPAT2, DGAT2, CIDE1, and TUSC5) was more extreme than that obtained after randomly sampling 1 million sets of the same number of genes (in all cases P ≤ 1.6 × 10^{-4}). The 12 genes in the TAG gene set exhibited a consistently lower expression in NSW than in WA (1.4-fold, P = 0.0007), independent of HGP treatment (P = 0.4; Fig. 5E). This is consistent with the greater IMF% observed in NSW animals in the full study (Cafe et al., 2011). The expression of the 12 TAG gene set genes in HGP-treated samples was reduced by an average of 1.6-fold (P = 0.0004). This effect was consistent across sites: 1.6-fold in WA (P = 0.0008) and 1.5-fold (P = 0.0003) in NSW (Fig. 5A). This result is partly consistent with the analysis of the impact of HGP on IMF% from the larger set of 255 animals in this experiment, where there was an impact of HGP on IMF% (P = 0.008) in animals on the WA site, reduction from 2.3% to 1.8% mean IMF (Cafe et al., 2010b). However, surprisingly, no effect of HGP treatment on IMF% (P = 0.29) was observed in the NSW animals (Cafe et al., 2010b).

Similarly, the tenderness genotype of the animals had no impact (P > 0.05) on the effect of HGP treatment on the expression of the 12 genes in the TAG gene set [tender Genotype 1.3-fold reduction (P = 0.007); tough Genotype 1.6-fold reduction (P = 0.0008); Fig. 5B]. Cattle possessing the favored tender alleles of CAPN
Figure 5. Differential expression of genes in the LM of Brahman cattle for various contrasts. The positions of the 12 triacylglyceride (TAG) synthesis and storage gene set genes used for this analysis (cell-death-inducing DNA fragmentation factor (DFFA)-like effector a (CIDEA), adipogenin (ADIG), S100 calcium binding protein G (S100G), phosphoenolpyruvate carboxykinase 1(PCK1), perilipin 1 (PLIN1), fatty acid binding protein 4, adipocyte (FABP4), adiponectin (ADIPOQ), plastin 1 (PLS1), 1-acylglycerol-3-phosphate O-acyltransferase 2 (AGPAT2), diacylglycerol O-acyltransferase homolog 2 (DGAT2), cell-death-inducing DFFA-like effector c (CIDEc), tumor suppressor candidate 5 (TUSC5)) are highlighted by black dots. P-values for the observed distributions of the 12 TAG gene set genes are as follows: (A) $7 \times 10^{-6}$, (B) $1.4 \times 10^{-4}$, (C) $<1 \times 10^{-6}$, (D) $2 \times 10^{-6}$, (E) $2.6 \times 10^{-5}$, and (F) $1.6 \times 10^{-4}$. NSW = New South Wales; WA = Western Australia; HGP = hormone growth promotant.
and CAST exhibited, on average, a 1.7-fold \( (P = 0.0003) \) reduction in the expression of the 12 genes in the TAG gene set compared with the animals possessing the tough alleles (Fig. 5C). This effect was larger in NSW (1.8-fold, \( P = 0.0002 \)) compared with WA (1.2-fold, \( P = 0.03 \)), and there was a difference \( (P = 0.008) \) between the 2 sites on the impact on gene expression of the 12 genes in the TAG gene set. Again, these results are partly consistent with analysis of the full experiment using 255 animals, of which these animals were a subset, where the favored \( CAPN3 + CAST \) alleles reduced mean IMF% from 2% to 1.7% \( (P = 0.05) \) for animals at the NSW site (Cafe et al., 2010b). No interaction between tenderness genotype and IMF% \( (P = 0.98) \) was detected in animals from WA in the larger data set (Cafe et al., 2010b). The animals possessing the tender genotype were associated with an average 1.2-fold decrease in the expression of the 12 TAG gene set genes in NSW relative to WA, whereas the animals possessing the tough genotype were associated with an average 1.2-fold increased expression in NSW relative to WA (Fig. 5F), the difference in effects between the 2 genotypes and states being highly significant \( (P = 0.0001) \). The HGP treatment status had only a small impact \( (P = 0.03) \) on the differences in expression of the 12 genes in the TAG gene set between tenderness genotypes (Fig. 5D).

**DISCUSSION**

**IMF Synthesis Pathways and Gene Expression Sets in Cattle Skeletal Muscle**

The analysis of the gene expression patterns in the high (WxH) and low (PxH) marbling crosses has identified a number of sets of genes likely to be involved in lipid metabolism. Consistent with previous work (Wang et al., 2005; Bonnet et al., 2007; Lee et al., 2007; Duckett et al., 2009; Graugnard et al., 2009; Wang et al., 2009b; Graugnard et al., 2010; Pickworth et al., 2011; Jeong et al., 2012), but more comprehensively, the expression of genes encoding proteins with roles in FA and TAG synthesis and storage appears to be correlated with a phenotype that measures the storage of TAG. The TAG gene set has high functional integrity, and all of the genes (except perhaps PLS1) were supported by the correlation of their expression with IMF% in the Brahman. The genes in the FA gene set were less supported by the correlation with IMF% and integrated analyses (12 of 25), but many of the highly supported genes encode products that have a clear link to FA synthesis. The genes in the PPARG gene set were much less well supported by the correlation with IMF% and integrated analyses (15 of 43) and appear to have less functional integrity. Overall, the results suggest that our detailed approach to the definition of gene sets is appropriate in this specific experimental context. The assignment of genes encoding proteins involved in TAG and FA synthesis to different gene sets on the basis of their expression profiles suggests that TAG and FA synthesis are regulated slightly differently across development. Interestingly, neither of the 2 regulators included in the genes sets, PPARG and CEBPA, are in the TAG or FA gene sets. The overall greater correlation with IMF% of genes in the TAG gene set is consistent with all paths to IMF requiring the synthesis and storage of TAG in the intramuscular adipocytes but not all paths requiring synthesis of FA in the intramuscular adipocytes.

However, if we had undertaken this analysis on just the animals from the NSW experimental site that were not treated with HGP, we would have come to a different conclusion (i.e., that there is no correlation between IMF% and the expression of genes encoding proteins involved in FA and TAG synthesis and storage). The difference in correlation at the 2 sites between gene expression and IMF% in untreated animals but not in the HGP-treated animals suggests that the difference is not due solely to the HGP treatment or to the experimental site. This unexpected result suggesting that under certain circumstances, yet to be accurately defined, the expression of the TAG genes is not an accurate indicator of IMF% requires further investigation.

**New Genes Encoding Productions with Potential Roles in IMF Deposition**

Although many of the genes in the TAG gene set encode products with clear roles in TAG synthesis and storage, the role of the products of a number of the other genes in the process is not so clear. Thyroid hormone responsive protein (THRSP; Cunningham et al., 1998), ADIPOQ (Shehzad et al., 2012), and ADIG (Hong et al., 2005) are regulatory/signaling proteins associated with lipid metabolism and thus may be involved in the regulation of TAG deposition or the signaling of TAG deposition in the intramuscular adipocytes to other parts of the system. Tumor suppressor candidate 5 (TUSC5; Oort et al., 2007; Sallman Almen et al., 2012) may also be a signaling protein. The potential roles of S100G and PLS1 in lipid metabolism are unknown. The composition of the FA gene set is more diverse than the TAG gene set and, outside of the genes encoding proteins with known roles in FA synthesis, contains a number of enzyme-coding genes and 2 genes encoding proteins secreted by adipocytes, \( CLU \) and \( RBP4 \) (Chen et al., 2005). The composition of the PPARG gene set is even more diverse; however, there is strong support for many of the genes from all of the analyses. For example, our results suggest that the iron-transporting protein transferrin (TF) may have a much
more important role in the intramuscular adipocytes than in the muscle contractile cells. Indeed, TF has been shown to be secreted by adipocytes (Zhou et al., 2009).

**Poorly Supported Genes Encoding Products with Key Roles in IMF Deposition or Correlated with IMF%**

A number of genes encoding products with key roles in TAG and FA synthesis or storage [e.g., members of the lipin (LPIN) gene family; Harris and Finck, 2011] were not strongly supported by this analysis. It is possible that the probes on the microarray for these genes are not returning reliable data, albeit many of these probes do show significant changes in signal across the development time course. It is also possible that these genes encode proteins with multiple functions in muscle within a single cell type or in the different cell types sampled or are active at different times in the different cell types. The expression of these genes would be subject to several different control mechanisms; hence, they would not be well correlated with a single phenotype.

For the LPIN gene family, LPIN1 was reported to account for all of the phosphatidate phosphatase type-1 activity in skeletal muscle (Donkor et al., 2007). Consistent with this, expression of LPIN3 was reported to be low or background in mouse and human muscle (Donkor et al., 2007) and is very low at all time points in the WxH and PxH development time series. The expression of LPIN2 is greatest at 60 d after conception, declining to a low and stable expression postnatally (again consistent with low or background expression in mouse and human muscles; Donkor et al., 2007), with no difference between PxH and WxH animals, consistent with no correlation of expression of LPIN2 with IMF% in the Brahman data set. The expression of LPIN1 also exhibited a large dynamic range in the developmental time course. Greatest expression was at birth, with much lower and stable expression after 7 mo, with no differences between the WxH and PxH crosses. The very different pattern of expression during development and lack of difference between WxH and PxH animals is again consistent with the lack of correlation of expression of LPIN1 with IMF% in the Brahman data set. The LPIN1 has dual functions, both enzymatic and regulatory (Harris and Finck, 2011). Therefore, it may be under a different regulatory program from the other TAG and FA synthesis and storage genes. In addition, in contrast to LPIN1, GPAM and HSB17D12, for example, are fairly well supported by the correlation and integrated analyses. Although they have distinctly different gene expression profiles across development from the genes encoding other members of the same biological process, the profiles are more similar to the TAG, FA, and PPARG gene sets gene expression profiles than the LPIN1 profile.

Of the genes encoding proteins with trans-2,3-enoyl-CoA reductase activity, TECR is the most likely candidate as TECR is an endoplasmic reticulum enzyme, whereas MECR and PECR encode enzymes with mitochondrial and peroxisomal locations, respectively. The TECR gene has a significant variation in expression in the LM development time course but has a profile different from any members of the TAG, FA, and PPARG gene sets and GPAM, HSB17D12, PTPLB, and LPIN1. However, the expression profile of TECR across the LM development time course is similar to the expression profile of PECR, which, unlike TECR, was located outside of the bulk of the distribution in the integrated analysis. There are few published data on TECR or its product, and the reason for the lack of correlation is obscure. Thus, for a variety of reasons, this analysis has not clearly identified the genes encoding the enzymes in intramuscular adipocytes responsible for the last step in conversion of C16 lipids to C18 lipids or in the synthesis of 1,2 diacyl-sn-glycerol from 1,2-diacyl-sn-glycerol 3-phosphate in the pathway to TAG.

The previously reported genes correlated with IMF% or marbling not supported by this analysis were not supported for a variety of reasons: possibly no signal from microarray (LEP and PDHB); expression signal observed, but very little variation in expression between samples (TGFβ1, RXRA, FAS, and SFRS18); variable expression level between samples, but the expression signal was not sufficiently greater in WxH than in PxH animals to enable the gene to pass the differential expression threshold for positively correlated genes or vice versa for negatively correlated genes (ADAMTS4, DLK1, RUNX1T1, CD36, SLC27A1, TTN, NEB, CAMK2A, HSPB1, AGT, PNPLA2, ACADVL, ACADM, EEF1A2, FABP3, MGP, OBSCN, TRDN, and RYR1); or sufficiently differentially expressed between WxH and PxH to pass the differential expression threshold (MGLL and LDLR), but correlation of gene expression with IMF% was not sufficiently high to pass the correlation threshold. The second category may reflect a high background signal masking true variation in expression or, alternatively, a very stably expressed gene. Thus, the correlation of muscle expression of LEP and PDHB, and TGFβ1, RXRA, FAS, and SFRS18 with IMF% or marbling may not have been adequately tested with the data sets described herein. However, clearly, most genes show variable expression (evidence that the probe is likely reporting bona fide gene expression) and are not supported either by the high vs. low IMF% comparison or by the correlation of gene expression with IMF%.

In addition, even the greatest correlations of gene expression with IMF% were low compared with other reported analyses, and the discrimination of less well
correlated genes from the background may have been substantially reduced in our analysis.

It is possible that level of IMF% had an impact, but our data sets covered a similar range of IMF% values to the other cattle and the pig data sets. We included 2 different diets and a diversity of cattle breeds, including both *B. taurus* and *B. indicus*; thus, it is biased toward genes encoding proteins with roles under a broader range of conditions. On the other hand, our analysis tests a large proportion of the protein-coding genes expressed in muscle, whereas many of the other studies considered only a much smaller number of candidate genes. In the latter case, it is much harder to evaluate the significance of the observed correlation coefficients relative to the unknown full distribution of correlation coefficients.

**Evaluating the TAG Gene Set as a Gene Expression Phenotype**

We demonstrated previously that steroid treatment of cattle reduces the expression of genes in LM believed to be involved in the storage of lipid in intramuscular adipocytes (De Jager et al., 2011). Here we show that this effect is consistent across both the experimental sites. However, this finding is not consistent with the phenotypic changes previously reported (for the larger data set), where it was shown that the steroid treatment reduced marbling in WA, an effect not observed in NSW (Cafe et al., 2010b). The observation on the larger data set for the NSW Brahman cattle is also at odds with the general observation that treatment with trenbolone acetate and estradiol leads to a reduced IMF% compared with untreated animals (Reiling and Johnson, 2003; Bryant et al., 2010; Hunter, 2010). This suggests that the gene expression phenotype may be more sensitive than the conventional approach or that the relationship between the expression of the genes and IMF% is more complex than it first appears.

The observation that the expression of the TAG gene set genes was not correlated with IMF% in untreated animals in NSW but was correlated in untreated animals in WA also suggested that differences between the sites has led to different relationships between expression of the TAG gene set genes and IMF%. It is plausible that the balance between the contributions to IMF% of past nutritional history and of the nutritional regimen at the time of sampling may have an impact on the extent of the correlation between TAG gene expression and IMF% at the time of sampling. In this experiment, the animals at the NSW experimental site had a larger ADG than the animals at the WA experimental site in the backgrounding phase and vice versa in the feedlot stage of growth (Cafe et al., 2010a). This is consistent with the feedlot stage having a greater impact on the gene expression phenotype observed at the time of muscle biopsy in the animals at the WA experimental site than the animals at the NSW experimental site.

It is generally believed that the favored CAPN/CAST tenderness markers do not reduce IMF%. Thus, our observation of an apparent negative interaction between the favored tenderness genotype and expression of the TAG gene set genes in animals from NSW but not in the animals from WA is unexpected. However, it is consistent with the observation from a larger number of animals from the same experiment that the cattle with the favored tenderness genotype for the *CAST* plus *CAPN3* genes in the NSW site had a small, but significant, reduction in IMF% (Cafe et al., 2010b). This result is also consistent with a previous study where it was found that the favorable *CAST* genotype was associated with leaner cattle (Wolcott and Johnston, 2009). However, the effect that we have observed is opposite to a recent study where the favorable allele of *CAPN1* appeared to be associated with increased marbling (Barendse, 2011). Animals with the favorable *CAST* genotypes previously have been found to have a tendency toward lower feed intake and more rapid growth in both cattle (Goll et al., 1998) and pigs (Kuryl et al., 2003), both traits likely to contribute to reduced IMF%.

In conclusion, the integrated analysis of gene expression data from 2 independent experiments in cattle identified the key output pathways for IMF% in LM and probably skeletal muscle in general. It also generated an informative tool for the prediction of interaction between the environment, genetics, and IMF%, which has been shown to be of utility using small numbers of animals, even with the very low IMF% typical of *B. indicus*. We propose that the use of the TAG gene set genes as an expression phenotype may represent a cost-effective alternative to traditional methods for the assessment of the impact of genetics and environment on IMF%.

**LITERATURE CITED**


De Jager et al.


