Distinctive gene expression patterns and imprinting signatures revealed in reciprocal crosses between cattle sub-species

Ruijie Liu 1, Rick Tearle 1, Wai Yee Low 1, Tong Chen 1, Dana Thomsen 1,3, Timothy P. L. Smith 2, Stefan Hiendleder 1,3 and John L. Williams 1,4*

Abstract

Background: There are two genetically distinct subspecies of cattle, Bos taurus taurus and Bos taurus indicus, which arose from independent domestication events. The two types of cattle show substantial phenotypic differences, some of which emerge during fetal development and are reflected in birth outcomes, including birth weight. We explored gene expression profiles in the placenta and four fetal tissues at mid-gestation from one taurine (Bos taurus taurus; Angus) and one indicine (Bos taurus indicus; Brahman) breed and their reciprocal crosses.

Results: In total 120 samples were analysed from a pure taurine breed, an indicine breed and their reciprocal cross fetuses, which identified 6456 differentially expressed genes (DEGs) between the two pure breeds in at least one fetal tissue of which 110 genes were differentially expressed in all five tissues examined. DEGs shared across tissues were enriched for pathways related to immune and stress response functions. Only the liver had a substantial number of DEGs when reciprocal crossed were compared among which 310 DEGs were found to be in common with DEGs identified between purebred livers; these DEGs were significantly enriched for metabolic process GO terms. Analysis of DEGs across purebred and crossbred tissues suggested an additive expression pattern for most genes, where both paternal and maternal alleles contributed to variation in gene expression levels. However, expression of 5% of DEGs in each tissue was consistent with parent of origin effects, with both paternal and maternal dominance effects identified.

Conclusions: These data identify candidate genes potentially driving the tissue-specific differences between these taurine and indicine breeds and provide a biological insight into parental genome effects underlying phenotypic differences in bovine fetal development.

Keywords: Cattle, Fetal development, Transcriptome
Background
There are substantial phenotypic and genetic differences among cattle breeds, in particular between indicine and taurine breeds (Bovine HapMap Consortium 2009). The taurine and indicine subspecies of cattle arose from independent domestication events resulting in a high degree of genetic divergence [1]. Phenotypically, indicine cattle are more tolerant of hot, humid environments and show greater resistance to parasites such as ticks; hence they are better adapted to survive in tropical areas [2]. However, the productivity of indicine cattle is lower than taurine cattle across a range of traits when measured in temperate zones, including growth and meat quality. Crossbreeding has been used to harness the positive traits of the two types to improve the performance of cattle in tropical environments [3]. Genes such as MSRB3 and PLAG1, which are involved in energy and muscle metabolism, have been shown to have subspecies-specific alleles that affect weight and body condition [4]. However, the genetic factors involved in adaptation to tropical conditions remain largely unknown.

Phenotypic differences between indicine and taurine breeds emerge during fetal development [5] and are reflected in birth outcomes, including birth weight [6]. Fetal growth rate accelerates after mid-gestation (~day 150) [7] and subspecies-specific phenotypes emerge. For example, taurine cattle have a greater myotube cross sectional area and greater bone size than indicine cattle at day 153 [8, 9]. Maternally inherited genes have been shown to contribute disproportionately to myofiber development and muscle and bone in reciprocal crosses, suggesting parent-of-origin imprinting effects [8, 9].

Advances in genome sequencing technology have facilitated the detailed exploration of transcriptome complexity and dynamics. Studies of gene expression in adult bovine tissues, including muscle [10], liver [11, 12], mammary gland [13] and adipose tissue [14] from either taurine or indicine breeds have identified genetic variation associated with differences in feed efficiency, milk composition and deposition of intramuscular fat. However, there is little information available on differences in gene expression between breeds during fetal development. A comparison of gene expression between taurine and indicine breeds may provide biological insights into the origin of their phenotypic differences.

This study investigated the transcriptome of the placenta and four somatic tissues at mid-gestation from two cattle breeds (Angus and Brahman) and their reciprocal crosses. The differentially expressed genes (DEGs) detected between the breeds and between the reciprocal crosses at this fetal stage represent candidates that may be involved in establishing phenotypic differences between the cattle subspecies.

Results
Expression profiles of five tissues
A total of 120 samples were analysed, which comprised brain, liver, lung, muscle and placenta samples from 3 pure Angus, 3 pure Brahman, 3 Brahman cross Angus and 3 Angus cross Brahman fetuses. Between 60 and 100 M 100 bp PE reads, or 90-130 M 75 bp PE reads per sample passed quality control. Reads were aligned to the extended Brahman reference genome (UOA_brahman_1 plus non-PAR Y chromosome from UOA_angus_1) using hisat2 with default settings, giving an average mapping rate of 89%. The total number of expressed genes among samples ranged from 16,368 to 17,013 and showed no substantial variation between tissues. There was a high correlation coefficient between expression of the same genes in each tissue in pure bred Brahman (Bi) and Angus (Bt) (Supplementary Fig. 1a-e). There were 14,143 genes expressed in all tissues (Supplementary Fig. 1f) with 5 genes consistently represented among the 20 most abundant transcripts in all five tissues: Insulin-Like Growth Factor 2 (IGF2), Eukaryotic Translation Elongation Factor 1 Alpha 1 (EEF1A1), Collagen Type III Alpha 1 Chain (COL3A1), Actin Beta (ACTB) and the paternally expressed gene 3 (PEG3).

Multi-scaling analysis grouped samples from each of the 5 tissues into tight clusters which were distinct from each other (Fig. 1a). A multi-factor model was used to account for and remove tissue effects, after which a PCA separated the samples by genetic groups in the first principle component (x-axis) and by sex in the second principle component (y-axis) (Fig. 1b). The expression for each tissue from each genetic type showed the same pattern within sex, with the 2 purebred groups well separated for all tissues, while the reciprocal crosses were less well separated (Supplementary Fig. 2a-e). The 20 most highly expressed genes in each tissue are reported in Supplementary Table 1.

Differential gene expression between purebred groups
There were 1085, 1495, 1935, 2515 and 2645 genes for which the normalized average number of mapped reads (CPM) differed significantly between purebred Bt and Bi brain, placenta, lung, liver and muscle, respectively. We designated these as differentially expressed genes (DEGs). Muscle had the largest number of DEGs among the tissues studied, but about 84% of these showed a fold change (FC) < 2, while in other tissues ~ 62–72% showed a FC < 2. The most significantly enriched gene ontology (GO) biological process and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in muscle included collagen metabolic process (GO:0032963); colla- gen fibril organization (GO:0030199); amino sugar and nucleotide sugar metabolism (bta00520) and glycin, serine and threonine metabolism (bta00260). Genes in
all four of these pathways had higher expression in Bt than in Bi.

Among the DEGs, ~10% in each tissue were IncRNAs. About 92% of DE IncRNAs had the opposite transcriptional direction to differentially expressed genes located within 100 kb.

**DEGs common to all five tissues in pure-bred groups**

There were 110 DEGs between Bi and Bt in common for all five tissues, comprising 50 annotated protein-coding genes, 42 genes lacking annotation in the reference genome and 18 IncRNAs (Fig. 2a). Alignment of the unannotated protein-coding genes to known genes in other cattle and ruminant reference genomes facilitated the annotation of 37 of the unnamed DEGs, based on >90% sequence identity. Of the 87 genes for which annotation was obtained (See Supplementary Table 2) and that were DE in all five tissues between the purebred animals, 84 had consistent relative abundance between subspecies Bt and Bi with respect to genotype in all tissues. The 3 exceptions were Aldehyde Oxidase 1 (AOX1), Choline Dehydrogenase (CHDH), Syntaxin 11 (STX11), whose expression was in a different direction (Bt vs Bi) in the liver compared with the other 4 tissues. GO pathway analysis of the set of 87 annotated genes showed that they were significantly enriched in 10 GO terms with \( p \)-value < 0.05, including

---

**Fig. 1** Multi-dimensional scaling (MDS) plot of sample expression profiles in five tissues. **a** The first two dimensions separate the samples by tissue type. **b** After accounting for the tissue source, samples are separated by genetic group in the first dimension (X-axis) and by sex in the second dimension (Y-axis). (1-pure Bt, 2-BtXBi, 3-BiXBt, 4-pure Bi. Male samples are shown in blue and female red)

---

**Fig. 2** DEG across 5 tissues. **a** Venn diagram depicting the distribution of DEGs across five tissues at FDR cut off 0.05. **b** Significantly enriched gene ontology terms for biological process (purple), Molecular function (red) and cellular component (blue) for 87 annotated DEGs genes that were in common across all five tissues. Bars indicate the percentage of DEGs in the GO term
oxidation-reduction process (GO:0055114), intracellular protein transport (GO:0006886), glycogen catabolic process (GO:0005980), positive regulation of protein autophosphorylation (GO:0031954) (Fig. 2b).

**Tissue-specific genes between purebred groups**

Genes that were DE between purebred Bt and Bi in only one of the five tissues examined were considered tissue specific DEGs. Using an FDR cut-off of $< 0.05$ and FC $\geq 2$, brain, liver, lung, muscle and placenta had 187, 328, 289, 388 and 191 tissue-specific DEGs respectively. GO biological process pathway enrichment analysis for these filtered tissue-specific DEGs identified 54 GO terms (Supplementary Table 3). The liver-specific DEGs were enriched for 6 GO terms including ion binding (GO: 0043167) and primary metabolic processes (GO: 0044238). Muscle was enriched for 9 GO terms including the collagen fibril organization pathway (GO: 0030199). Brain was also enriched for 9 GO terms that included pathways involved with detection of stimulus (GO:0050906) and nervous system processes (GO:0050087). Lung was enriched for 10 GO terms, most of which were related to fundamental biological processes, including regulation of molecular function (GO:0065009) and cellular response to endogenous stimulus (GO:0071495). Placenta was enriched for 20 GO terms which were linked to proton-transporting V-type ATPase (GO:0033176) and domain small molecule metabolic process (GO 0044281).

**Differential gene expression between crossbred groups**

Comparison of transcript abundance between the reciprocal cross-bred groups (Bt x Bi and Bi x Bt) did not reveal a substantial number of DEGs ($<20$/tissue at FDR $< 0.05$), except for liver which had 2473 DEGs. However, only 143 (5.8%) of the liver DEGs had a fold change greater than 2. We performed GO biological process pathway enrichment analysis and KEGG pathway enrichment analysis for the protein coding DE genes with $>2$-fold change. The GO analysis showed that DEGs were significantly enriched in 6 GO terms, including: macromolecule metabolic process (GO:0043170), primary metabolic process (GO:0044238), cellular metabolic process (GO:0044237), metabolic process (GO:0006807), and organic substance metabolic process (GO:0071704) which are all involved in metabolic processes. The only significantly enriched KEGG pathway was metabolic pathways (path: bta01100).

Pairwise comparisons of the DEGs in liver for the 4 genetic groups were performed to explore relationships in expression patterns between pure bred and crossbred concepti. The sire dominated the liver expression pattern in Bt-sired crossbred (Bt x Bi) liver which had 1276 DEGs when compared to purebred Bi liver, versus 219 DEG when compared with purebred Bt liver. However, the dam breed appears to dominate expression pattern in Bi-sired crossbreds, with 317 DEGs in the Bi x Bt crossbred compared with purebred Bt, but 150 DEGs when compared with purebred Bi liver transcripts.

**Expression pattern of DEGs from the purebred groups in comparison with crossbred groups**

The expression pattern of the 6456 DEGs between tissues of purebred animals was examined in the reciprocal crossbred groups. Of these DEGs 5784 (~ 90%) showed an additive expression pattern where both paternal and maternal genomes contributed to the gene expression levels in the crossbred groups (Fig. 3a), as suggested by the transcript abundance falling approximately midway between that of the two purebred classes. However, transcript abundance of some DEGs (672) was more consistent with parent-of-origin driven expression (Fig. 3b-i). Different types of such effects were observed, predominantly maternal/paternal dominance and Bt or Bi allele derived dominance. The abundance of DEGs between crossbred groups fell into three general categories: co-dominant, dominant and recessive expression patterns, with dominance in some cases driven by either the male or the female (Fig. 3). The number of genes falling into each category are given in Table 1.

GO analysis of the DEGs that overlapped between tissues showed that they were significantly enriched in 19 GO terms including positive regulation of cellular metabolic process (GO:009893), positive regulation of nitrogen compound metabolic process (GO:0051773) and membrane-enclosed lumen (GO:0031974). The transcript levels of the DEGs involved in these significantly enriched pathways had exclusively higher expression in the purebred Bt compared with the purebred Bi.

**Discussion**

The study of gene expression in prenatal development will help to understand the regulation of fetal tissue-specific growth and development. Our hypothesis was that phenotypic differences between subspecies of cattle may be due, in part, to differential gene expression during mid-gestation. Consistent with this hypothesis, in this study we observed substantial differences in expression between breeds of cattle from the two genetically distinct sub-species *Bos taurus taurus* (Angus) and *Bos taurus indicus* (Brahman). In addition, we observed differential expression of genes in reciprocal crosses between these subspecies, some of which revealed parent-of-origin and breed-of-origin effects on gene expression in five tissues at mid-gestation.
We found that five genes had high levels of expression in all five tissues at this developmental stage (IGF2, EEF1A1, COL3A1, ACTB, and PEG3). These genes play a crucial role in embryonic development and fetal tissue growth, as shown by loss-of-function mutations which result in developmental delay and several diseases including intellectual disability, immune system abnormalities, cerebral abnormalities and abnormally large abdominal organs [15–19]. EEF1A1 is a member of the eukaryotic elongation factor family that regulates protein synthesis, that is expressed in brain, placenta, lung, liver, kidney, and pancreas in human adults [20]. COL3A1 is expressed in extensible connective tissues, such as skin and lung. A mutation in COL3A1 has been linked to vascular disease [21]. Expression levels of IGF2 have been linked to increased muscle mass [22] and fetal growth [23].

Other highly abundant transcripts showed tissuespecific expression levels which were related to tissue function. Alpha-Fetoprotein (AFP) had liver-specific expression and encodes a major plasma protein produced by the liver during fetal development [24]. Two genes that were highly expressed in the muscle were the muscle structural protein genes Myosin Heavy Chain 3 (MYH3) and Myosin Binding Protein C, Slow Type (MYBPC1) [25, 26]. Genes that play an important role in neurodevelopment including Adenylate Cyclase 1

| Table 1 Number of genes showing a parent of origin effect on expression patterns in five tissues |
|---------------------------------|-------------|-------------|-------------|-------------|-------------|
|                                | Brain       | Liver       | Lung        | Muscle      | Placenta    |
| Maternal genome driven - Taurine| 0           | 12          | 2           | 16          | 0           |
| Maternal genome driven - Indicine| 0           | 11          | 4           | 11          | 5           |
| Paternal genome driven – Taurine| 6           | 24          | 9           | 16          | 7           |
| Paternal genome driven - Indicine| 5           | 36          | 13          | 13          | 12          |
| Taurine dominant – activation  | 27          | 65          | 40          | 33          | 25          |
| Taurine dominant – inhibition  | 7           | 30          | 24          | 10          | 12          |
| Indicine dominant – activation | 13          | 38          | 17          | 15          | 10          |
| Indicine dominant - inhibition | 2           | 22          | 3           | 15          | 5           |
(ADCY7), Statmin 2 (STMN2) and Tubulin Beta 3 Class III (TUBB3) were highly expressed and specific to the brain [27–29]. All of these genes had high levels of expression in both the pure breed concepti and the crosses. The lung was the only tissue that did not have any highly expressed tissue-specific genes (cut off Log2CPM > 10) at this developmental stage.

Muscle composition and quality of taurine and indicine cattle breeds differs [30] and is largely determined during fetal development [31]. We have previously reported greater fast myotube cross sectional area and greater bone size in taurine than indicine cattle fetuses at day 153 [8, 9]. In the present study we show that muscle contained the highest number of DE genes between purebreds amongst all studied tissues. Significantly enriched pathways included collagen metabolic process, collagen fibril organization, amino sugar and nucleotide sugar metabolism and glycerine, serine and threonine metabolism. Genes in all four of these pathways had higher expression in Bt than in Bi fetuses. Although we did not examine gene expression in bone tissue, it is known that fetal muscle and bone growth are linked and collagen pathways also play a major role in bone growth [32].

Intrauterine stress increases the risk of adult disease through fetal programming mechanisms. Increased oxidative stress during embryonic and fetal growth can be caused by environmental and physiological conditions [33], and may affect key transcription factors that can alter gene expression during development [34]. From the GO pathway analysis in the current study, oxidation-reduction processes and oxidoreductase activity were found to be significantly associated with the DEGs between the two pure breeds that were in common to all five tissues.

Heat shock leads to oxidative stress, which has been associated with reduced production performance in Bos taurus indicus [35]. During heat stress the steady-state concentration of free radicals is disturbed, resulting in both cellular and mitochondrial oxidative damage [36]. A study of the effects of oxidative stress on cattle fertility indicated that in tropical areas, Bos taurus taurus bulls have a higher level of reactive oxygen species (ROS) in their semen than Bos taurus indicus bulls [37]. It has been suggested that these high levels of ROS cause major sperm defects in heat stressed Bos taurus taurus bulls [34]. In our study, TXNRD2, a nuclear genome encoded mitochondrial protein that scavenges reactive oxygen species, had a higher level of expression in Bi than Bt in all tissues. It is possible that TXNRD2 mediated protection of mitochondrial function may help indicine cattle to better adapt to hot environments.

The HSD11B1L encoded protein catalyses the interconversion of inactive to active glucocorticoids, e.g. the conversion of inactive cortisone to the active forms: corticosterone and cortisol. These are key hormones that regulate a variety of physiologic responses to stress through the hypothalamus-pituitary-adrenal (HPA) axis that is responsible for the adaptation of stress responses to restore homeostasis [38]. Higher levels of HSD11B1L transcripts were found in all Bi tissues compared with Bt, which may allow indicus cattle to respond more rapidly than taurine cattle to stressful situations, including environmental and biological challenges.

Most of the genes that were DE in all five tissues showed changes in the level of expression in the same direction for all tissues. There were 3 exceptions with different directions of expression in the liver compared with the other 4 tissues. The liver plays an important role in metabolic processes and in immune system function, which affects the response to many diseases [39, 40]. We found that the expression of AOX1 was higher in all Bi tissues except liver, where it was lower. AOX1 produces hydrogen peroxide and catalyses the formation of superoxide. Levels of AOX1 increase in mouse liver following infection [41] suggesting a role in immune response by stimulating host immunity, inflammation and coagulation. Indicine cattle are generally less susceptible to disease than taurine cattle [42, 43]. For example, they are more resistant to ticks [44] and tuberculosis [45]. Interestingly AOX1 had lower levels of expression in Bi than Bt in tissues other than liver. The significance of this is unclear. The GO terms including genes that were DE between purebreds in this study showed that those involved in metabolic processes generally had significantly higher expression in Bt compared with Bi. Low metabolic rate has been associated with thermotolerance of Bos taurus indicus [46].

Interestingly, the genes that were DE between the liver of the pure-bred concepti, that were also differentially expressed between the reciprocal crossbred concepti, showed a higher expression when the sire was taurine for both sexes. For example, a critical nuclear receptor NR4A1 had a higher level of expression in pure Bt and in the crossbred concepti when the sire was Bt. NR4A1 is involved in inflammation, apoptosis, and glucose metabolism and also regulates a paternally imprinted gene, SNRPN, which affects neurological and spine development [47]. NR4A1 regulates energetic competence of mitochondria and promotes neuronal plasticity. However, studies in animal models and of neuropa-thologies in humans have shown that sustained expression of this gene results in increased sensitivity to chronic stress [48]. Higher levels of expression in Bt may be related to a reduced tolerance of stress including heat and drought conditions.

Genomic imprinting, which is reflected in a biased level of expression of one autosomal copy of a gene and is dependent on the parent of origin, has been reported
in all mammalian species in which it has been assessed, e.g. mice [49], humans [50], and domesticated animals [51]. Both insulin-like growth factor (IGF2) and paternally expressed gene 3 (PEG3) are imprinted in humans, mice, cattle and other species [52, 53] and the paternally inherited copy is expressed during fetal development, with expression declining rapidly after birth [54]. Both genes play an important role in controlling fetal growth rate and nurturing behaviours in mammals. In the present study, IGF2 and PEG3 were highly expressed in all samples across the 4 pure and crossbred groups in all five tissues, suggesting that both PEG3 and IGF2 functions are essential at mid-gestation. The overall levels of PEG3 and IGF2 transcripts did not differ between breeds or the direction of the cross, although we were unable to assign transcripts to a parent of origin to test for imprinting.

**Conclusion**

This study identified a large number of genes that showed significant tissue-specific expression differences between the taurine and the indicine breeds studied. These genes were found to participate in pathways related to tissue-specific function. Genes that were differentially expressed between Angus and Brahman in all tissues were found to relate to functions such as immune response and stress response, that may to some extent explain the higher resilience of Bi cattle. This study also identified genes that putatively have parent or breed of origin-controlled expression patterns. Exploring these further would require e.g. long read Iso-seq data to resolve haplotype specific expression. The current data provide a basis for future research on parental genome effects underlying phenotypic differences in cattle fetal development. Taking these factors into account in breeding and management may improve the welfare and productivity of cross-bred cattle in tropical environments.

**Material and methods**

**Animals and sample collection**

All animal experiments and procedures described in this study were compliant with national guidelines and approved by the University of Adelaide Animal Ethics Committee which follows ARRIVE Guidelines (https://arriveguidelines.org/) for approval and monitoring all studies involving live animals (Approval No. S-094-2005). The animals and semen used were pure bred Angus (Bos taurus taurus) and Brahman (Bos taurus indicus) cattle, subsequently referred to as Bt and Bi respectively. Purebred Bt and Bi females (heifers) of approximately 16–20 months of age were maintained on pasture supplemented with silage. The heifers were inseminated with semen of purebred Bt or Bi sires and pregnancy tested by ultrasound scanning. Pregnant heifers and their concepti were humanely sacrificed at day 153+/−1 of gestation and the conceptus dissected. Tissues were snap-frozen in liquid nitrogen and then stored at -80°C as previously described [8]. The five tissues used in this study, brain, liver, lung, muscle and placenta, were taken from 3 male and 3 female concepti, from each of the 4 genetic combinations (Bt x Bt, Bi x Bt, Bt x Bi, Bi x Bi; paternal genome listed first), giving a total of 24 samples per tissue.

**RNA isolation, library preparation and sequencing**

Total RNA was isolated from tissues using the RiboZero Gold kit, in accordance with the manufacturer’s recommendations (Illumina, San Diego, CA). Sequencing libraries were prepared with a KAPA Stranded RNA-Seq Library Preparation Kit following the Illumina paired-end library preparation protocol (Illumina, San Diego, CA). Paired-end (PE) sequence reads were produced on an Illumina NextSeq500 platform, 2 × 75 bp for placenta, lung and brain and 2x100bp for liver and muscle.

**Data analysis**

FastQC [55] was used to assess read quality and adaptor sequences were removed using cutadapt (Martin, 2011). The UOA_Brahman and UOA_Angus genome assemblies (GCA_003369695.2; GCA_003369685.2) are more contiguous that the ARS-UCD1.2 assembly and are completely phased, for this reason, and that data were produced from Brahman and Angus fetuses, these sequences were chosen as the reference. RNA seq reads were aligned with both UOA_Brahman and UOA_Angus assemblies and better alignment was found using UOA_Brahman. Approximately 93.1% sequences aligned to the Brahman genome whereas only 90.3% sequences aligned to the Angus genome. Therefore, an extended bovine Brahman reference genome, consisting of the autosomes and X chromosome from UOA_Brahman_1 and the non-PAR Y chromosome from UOA_Angus_1 was used in the analyses. Reads were aligned to this reference using hisat2 [56]. The number of annotated clean reads for each gene was calculated using feature counts from the Rsense and Ensembl annotation v97. Genes with a count per million (CPM) reads below 0.5 were excluded. Multi-dimensional scaling (MDS) plots were created using plotMDS from the limma R package. The expression of genes was normalised across the libraries by the Trimmed Mean of M-values (TMM) [58], and potential batch effects due to samples being sequenced in different sequencing runs were accounted for using the RemoveBatchEffect function in the limma package. Ignoring sex difference, differentially expressed genes (DEGs) with a false discovery rate (FDR) < 0.05 after down-weighting high variation replicates, were identified using the
**Functional analysis of DEGs**

To facilitate functional analysis of DEGs, cattle gene IDs were converted to homologous human Ensembl gene IDs using BioMart R packages [62]. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analyses of DEGs were performed using the limma R package [63]. GO terms for molecular functions (MF), biological processes (BP) and cellular components (CC) were interrogated. Fisher’s exact tests were carried out and an adjusted P value calculated using the Benjamini-Hochberg procedure for multiple tests (FDR). GO and KEGG terms with an adjusted P value < 0.05 were considered to be significantly enriched pathways. GSEA software was used to define and plot the pathway networks for DEGs.

**Identification of Brahman and Angus gene expression pattern in crossbred groups**

Some phenotypes of the concepti differed depending on which was the paternal breed in Bt and Bi reciprocal crosses, e.g. birth weight and mature size [8, 9]. We tested whether this difference was reflected in gene expression differences during fetal development. Genes where total abundance of transcripts was dominated by paternal breed were defined as paternally-driven, or when abundance was dominated by the maternal breed to be maternally-driven. To identify such genes, we used the difference in expression between purebred concepti to define breed-specific transcript abundance. Specifically, the absolute difference of transcript abundance in Bt versus Bi provided a “normalization” of expected abundance for a crossbred conceptus. Using that difference as denominator, we then used the absolute difference in abundance between the reciprocal crosses to identify genes influenced by the parental origin.

For example, if we assume a transcript has a difference in abundance between Angus and Brahman (Bt x Bt / Bi x Bi) of 3.0 (i.e. higher in Angus) and the difference in abundance in Angus-sired crossbreds compared with Brahman-sired crossbreds (Bt x Bi / Bi x Bt) is 2.7, the ratio of differences is 0.9 and is consistent with parental-driven expression. We use an arbitrary threshold of > 0.8 to designate parental-driven expression. We then use the differences in expression between the purebreds and crossbreds to determine if the gene is maternally or paternally driven. For convenience, we define the four genetic groups, i.e., the purebreds and crossbreds, by digits, with Bt x Bt, Bi x Bt, Bt x Bi, and Bi x Bi labelled as 1, 2, 3 and 4 respectively. All combinations of pairs of groups were compared, and for each gene the above ratio was calculated. Six combinations of expression differences were obtained: diff1–2, diff1–3, diff1–4, diff2–3, diff2–4 and diff3–4. The threshold was adjusted to identify genes in the following patterns:

1. Maternally driven-taurine: diff2–3/diff1–4 > 0.8, diff1–2/diff1–4 < 0.2, diff3–4/diff1–4 < 0.2, high expression level in Bt
2. Maternally driven-indicine: diff2–3/diff1–4 > 0.8, diff1–2/diff1–4 < 0.2, diff3–4/diff1–4 < 0.2, high expression level in Bi
3. Paternally driven-taurine: diff2–3/diff1–4 > 0.8, diff1–3/diff1–4 < 0.2, diff2–4/diff1–4 < 0.2, high expression level in Bt
4. Paternally driven-indicine: diff2–3/diff1–4 > 0.8, diff1–3/diff1–4 < 0.2, diff2–4/diff1–4 < 0.2, high expression level in Bi
5. Taurine dominant-Inhibition: diff2–3/diff1–4 < 0.2, diff24/ diff1–4 < 0.8, diff3–4/diff1–4 < 0.8, high expression level in Bi
6. Taurine dominant-Activation: diff2–3/diff1–4 < 0.2, diff24/ diff1–4 < 0.8, diff3–4/diff1–4 < 0.8, high expression level in Bt
7. Indicine dominant-Inhibition: diff2–3/diff1–4 < 0.2, diff1–2/ diff1–4 < 0.8, diff3–4/diff1–4 < 0.8, high expression level in Bt
8. Indicine dominant-Inhibition: diff2–3/diff1–4 < 0.2, diff1–2/ diff1–4 < 0.8, diff3–4/diff1–4 < 0.8, high expression level in Bi

Demonstrations of Bi and Bt gene expression patterns in crossbred groups are shown in Supplementary Fig. 3.

**Abbreviations**

Bi: Bos taurus taurus; Bt: Bos taurus indicus; DEG: differentially expressed genes; PE reads: paired end reads; MDS: multi-dimensional scaling; PCA: principal components analysis; CPM: counts per million.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s12864-021-07667-2.

**Acknowledgements**

We thank Kelsey McClure, Kristen Kuhn, William Thompson, and Bob Lee for technical support. Mention of trade names or commercial products in this publication is solely for information and does not imply recommendation or endorsement by USDA. USDA is an equal opportunity provider and employer.

**Authors’ contributions**

JLW, SH and TPLS conceived and managed the project. SH designed and obtained Bos taurus and Bos indicus fetal resources. DAT and TC extracted RNA samples and performed QC. TPLS produced sequence data. RL analysed
RNA seq data; RL, RT and WYL interpreted data. RL and JLW drafted the manuscript and all authors read, edited, and approved the final manuscript.

Funding
The study was funded from the JS Davies bequest through the Davies Research Centre, and by USDA-ARS Project 3040–31000–100-00-D.

Availability of data and materials
Sequence data is available from the GEO under accession number GSE148909.

Declarations
Ethics approval and consent to participate
All animal work was approved by the University of Adelaide Animal Ethics Committee, approval No. S-094-2005.

Consent for publication
Not Applicable

Competing interests
The authors declare no competing interests.

Author details
1Davies Research Centre, School of Animal and Veterinary Sciences, The University of Adelaide, Adelaide, Australia. 2USMARC, USDA-ARS-US Meat Animal Research Center, Clay Center, NE, USA. 3Robinson Research Institute, The University of Adelaide, Adelaide, Australia. *Present address: Dipartimento di Scienze Animali, della Nutrizione e degli Alimenti, Università Cattolica del Sacro Cuore, Piacenza, Italy.

Received: 21 December 2020 Accepted: 21 April 2021

Published online: 03 June 2021

References
1. Pitt D, Sevane N, Nicolazzi EL, MacHugh DE, Park SDE, Colli L, et al. Differences in beef quality between Angus (Bos taurus taurus) and Nellore (Bos taurus indicus) cattle through a proteomic and phosphoproteomic approach. PLoS One. 2017;12(1):e0170294. https://doi.org/10.1186/s12864-015-2292-8.
2. Mukibi R, Vinsky M, Keeogh KA, Fitzsimmons C, Stothard P, Waters SM, et al. Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. Sci Rep. 2018;8(1):1–12.
3. Cui X, Hou Y, Yang S, Xie Y, Zhang S, Zhang Y, et al. Transcriptional profiling of mammary gland in Holstein cows with extremely different milk protein and fat percentages using RNA sequencing. BMC Genomics. 2014;15(1):226. https://doi.org/10.1186/1471-2164-15-226.
4. Sheng X, Ni H, Liu Y, Li J, Zhang L, Guo Y. RNA-seq analysis of bovine intramuscular, subcutaneous and peripheral adipose tissues. Mol Biol Rep. 2014;41(3):1631–7. https://doi.org/10.1007/s12053-013-3010-8.
5. Curley JP, Barton S, Surani A, Keverne EB. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. Proc R Soc Lond B Biol Sci. 2004;271(1545):1305–9. https://doi.org/10.1098/rspb.2004.2725.
6. Abbas W, Kumar A, Herbein G. The eEf1A proteins: at the crossroads of oncogenesis, apoptosis, and viral infections. Front Oncol. 2015;5:75.
7. Horn D, Siebert E, Seidel J, Rost I, Mayer K, Abou Jamra R, et al. Biallelic COL3A1 mutations result in a clinical spectrum of specific structural bone anomalies and connective tissue abnormalities. Am J Med Genet A. 2017;173(9):2534–8. https://doi.org/10.1002/ajmg.a.38345.
8. Cuverton S, Stuart HM, Chandler KE, Roberts NA, Armstrong R, Bernardini L, et al. ACTB loss-of-function mutations result in a pleiotropic developmental disorder. Am J Hum Genet. 2017;101(6):1021–33. https://doi.org/10.1016/j.ajhg.2017.11.006.
9. Azzi S, Habil WA, Netchine I. Beckwith–Wiedemann and Russell–silver syndromes: from new molecular insights to the comprehension of imprinting regulation. Curr Opin Endocrinol Diab Obes. 2014;21(1):30–8. https://doi.org/10.1097/MED.000000000000037.
10. Harney JJ, Wilkins MR. Methylation of elongation factor 1A: where, who, and why? Trends Biochem Sci. 2018;43(3):211–23. https://doi.org/10.1016/j.tibs.2018.01.004.
11. Cortini F, Marinelli B, Romi S, Seresini A, Pesatori AC, Seia M, et al. A new COL3A1 mutation in Eilers-Danos syndrome vascular type with different phenotypes in the same family. Vasc Endovasc Surg. 2017;51(3):141–5. https://doi.org/10.1177/1538574417692114.
12. Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, et al. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature. 2003;238:32–6.
13. Kadakia R, Josefsen J. The relationship of insulin-like growth factor 2 to fetal growth and adipoysis. Horm Res Paediatr. 2016;85(2):75–82. https://doi.org/10.1159/000443500.
14. Pett FM, Hebert MB, Picone O, Brisset S, Maurin MA, Parisot F, et al. A new mutation in the AFGP gene responsible for a total absence of alpha fetoprotein on second trimester maternal serum screening for Down syndrome. Eur J Hum Genet. 2009;17(3):387–90. https://doi.org/10.1038/ejhg.2008.186.
15. Zieba J, Zhang W, Chong XJ, Forlenza KN, Martin JH, Heard K, et al. A postnatal role for embryonic myosin revealed by MYH3 mutations that alter TGFβ signaling and cause autosomal dominant spondyloepiostylosis. Sci Rep. 2017;7(1):41803. https://doi.org/10.1038/srep41803.
16. Ha K, Buchan IG, Alvarado DM, McCall K, Vidyarthi A, Luther PK, et al. MBP1C mutations impair skeletal muscle function in zebrafish models of arthrogryposis. Hum Mol Genet. 2013;22(24):4967–77. https://doi.org/10.1093/hmg/ddt344.
17. Wang H, Ferguson GD, Pineda WV, Cundiff PE, Storm DR. Overexpression of type-I adenylyl cyclase in mouse forebrain enhances recognition memory and LTP. Nat Neurosci. 2004;7(8):835–42. https://doi.org/10.1038/nn1248.
18. Wang Q, Zhang Y, Wang M, Song W-M, Shen Q, McKenzie A, et al. The landscape of multiscale transcriptomic networks and key regulators in Parkinson’s disease. Nat Commun. 2019;10:1–15.
19. Fukushima S, Kato M, Kawamura K, Tagaki T, Tsutsui A, Tsutsui H. A mutation in the tubulin-encoding TUBB3 gene causes complex congenital malformations and unilateral hypohydrosis. Child Neurol Open. 2016;3:2300481X16665758.
20. Rodrigues RT, Chizzotti ML, Vital CE, Baracat-Pereira MC, Barros E, Busato KC, et al. Differences in beef quality between Angus (Bos taurus taurus) and Nellore (Bos taurus indicus) cattle through a proteomic and phosphoproteomic approach. PLoS One. 2017;12(9):e0170294.
31. Picard B, Lefaucheur L, Berri C, Duclos M. Muscle fibre ontogenesis in farm animal species. Reprod Nutr Dev. 2002;42(5):415–31. https://doi.org/10.1051/rnd:2002035.

32. Salhotra A, Shah HN, Levi B, Longaker MT. Mechanisms of bone development and repair. Nat Rev Mol Cell Biol. 2020;21(1):696–711. https://doi.org/10.1038/s41580-020-00279-w.

33. Thompson LP, Al-Hassan Y. Impact of oxidative stress in fetal programming. J Pregnancy. 2012;2012:1–8. https://doi.org/10.1155/2012/582748.

34. Denenery PA. Role of redox in fetal development and neonatal diseases. Antioxid Redox Signal. 2004;6(1):1147–53. https://doi.org/10.1089/153208604771978453.

35. Fedyaeva A, Stepansov A, Lyubushkina I, Pobezhimova T, Rikhvanov E. Heat shock induces production of reactive oxygen species and increases inner mitochondrial membrane potential in winter wheat cells. Biochem Soc Trans. 2014;42(9):1202–10. https://doi.org/10.1042/BST20140127.

36. Belhadj Slimen I, Najari T, Ghram A, Abdrrabba M. Heat stress effects on livestock: molecular, cellular and metabolic aspects, a review. J Anim Physiol Anim Nutr. 2016;100(3):401–12. https://doi.org/10.1111/jpn.12379.

37. Nichi M, Boli P, Züge RM, Rambabe VH, Goovaerts I, Rambabe RC, et al. Seasonal variation in semen quality in Bos indicus and Bos taurus bulls raised under tropical conditions. Theriogenology. 2006;66(4):822–8. https://doi.org/10.1016/j.theriogenology.2006.01.056.

38. Walker JI, Spiga F, Gupta R, Zhao Z, Lightman S, Terry J. Rapid intra-adrenal feedback regulation of glucocorticoid synthesis. J Steroid Biochem. 2015;12(12):20140875. https://doi.org/10.1016/j.jsb.2014.0875.

39. MacParland SA, Liu MC, Mc Y-Z, Innes BT, Sartczak AM, Gage BK, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun. 2018;9(1):1–21.

40. Chen H, Meng H-Y, Liu S-M, Wang Y, Yang X-X, Lu F, et al. Identification of key metabolic changes during liver fibrosis progression in rats using a urine and serum metabolomics approach. Sci Rep. 2017;7(1):1–12.

41. Maeda K, Ohno T, Igarashi S, Yoshimura T, Yamashiro K, Sakai M. Aldehyde oxidase 1 gene is regulated by Nrf2 pathway. Gene. 2012;505(2):374–8. https://doi.org/10.1016/j.gene.2012.06.010.

42. Mackinnon M, Meyer K, Hetzel D. Genetic variation and covariation for growth, parasite resistance and heat tolerance in tropical cattle. Livest Prod Sci. 1991;27(2):310–22. https://doi.org/10.1016/0301-6226(91)90088-E.

43. Vajana E, Barbato M, Colli L, Milanesi M, Rochat E, Fabrizi E, et al. Combining landscape genomics and ecological modelling to investigate local adaptation of indigenous Ugandan cattle to East Coast fever. Front Genet. 2018;9:385. https://doi.org/10.3389/fgene.2018.00385.

44. Franzin AM, Maruyama SR, Garcia GR, Oliveira RP, Ribeiro JMC, Bishop R, et al. Estrogen and serum metabolomics approach. Sci Rep. 2017;7(1):1–12.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.