Proc. Assoc. Advmt. Anim. Breed. Genet. 23:63-66

PRE- AND POST-PUBERTY CO-EXPRESSION GENE NETWORKS FROM RNA-SEQUENCING OF BRAHMAN HEIFERS

L.T. Nguyen[1,2,3], A. Reverter[4], A. Cánovas[5], L.R. Porto-Neto[4], B. Venus[3], A. Islas-Trejo[6], S.A. Lehnert[4], J.F. Medrano[6], M.G. Thomas[7], S.S. Moore[3] and M.R.S. Fortes[1]

[1]School of Chemistry and Molecular Biosciences, The University of Queensland, Australia
[2]Faculty of Biotechnology, Vietnam National University of Agriculture, Vietnam
[3]Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Australia
[4]CSIRO Agriculture and Food, Queensland Bioscience Precinct, Australia
[5]Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, N1G 2W1, Canada
[6]Department of Animal Science, University of California-Davis, CA, USA
[7]Department of Animal Science, Colorado State University, Fort Collins, CO 80523, USA

SUMMARY

Brahman cattle, a *Bos indicus* breed, are well adapted to the harsh environment of northern Australia but reach puberty at an older age compared to *Bos taurus* breeds. Samples from hypothalamus (HYP), pituitary gland (PIT), both ovaries (OVA), liver (LIV), adipose tissue (AT), uterus (UTE) and *longissimus dorsi* muscle (MUS) from pre- and post-pubertal heifers were harvested for RNA sequencing (RNA-Seq). Four gene categories, including differentially expressed (DE) genes, tissue specific (TS) genes, key transcription factors (TF) and genes harbouring SNP associated with heifer fertility, were utilized as nodes of the gene co-expression networks. Significant network connections were identified using an algorithm that exploits the dual concepts of partial correlation and information theory (PCIT). Significance analysis ($P < 0.01$) of RNA-Seq data revealed 2,116 DE genes, 624 TS genes, 186 TF and 179 genes having SNP associated with heifer fertility within the 14,437 expressed genes (genes with reads per kilobase of exon per million mapped reads (RPKM) > 0.2). PCIT analysis pinpoints *ZEB1*, *TEF* and *NFATC2* as the best trio of TF in terms of their ability to span the majority of the topology of the pre- and post-puberty networks. A new role for *SEMA7A* in bovine pubertal development is also postulated. Taken together, our multi-tissue omics analysis revealed candidate genes that could lead to improved understanding of the mechanisms that guide pubertal development.

INTRODUCTION

Fertility traits are economically important for beef cattle operations. Improvements in reproductive efficiency can increase profitability and reproduction rate of beef cattle. Although events involved in the puberty process are similar in *Bos indicus* and *Bos taurus* cattle, they are initiated earlier in *Bos taurus* (Johnston *et al*., 2009). Selection programs for early pubertal cattle based on phenotype require additional expenditure and labour. As the precise mechanisms inhibiting or stimulating bovine puberty are not entirely clear, identification of molecular regulatory networks modulating puberty in *Bos indicus* cattle is required to better manage heifer development, support development of new biotechnologies, and perhaps develop genetic selection tools of early pubertal cattle.

Our study aimed to identify DE genes, TF, metabolic pathways and networks involved in Brahman cattle puberty. Key tissues for puberty (HYP, PIT, OVA and UTE) and for growth and metabolism (LIV, MUS and AT) were collected from six pre- and six post-pubertal Brahman heifers for RNA-Seq analyses. Gene expression values were obtained and used to construct pre- and post-puberty co-expression gene networks using an algorithm based on PCIT. The predicted co-expression networks were linked by DE genes, TS genes, known TF and genes harbouring SNP associated with

heifer fertility traits. These analyses provide new insights into candidate regulatory genes and gene expression pathways involved in bovine puberty.

MATERIALS AND METHODS

Twelve heifers of similar age were managed, handled and euthanized under protocols approved by the Animal Ethics Committee of the University of Queensland (UQ), Production and Companion Animal group (certificate number QAAFI/279/12). Heifers were examined every two weeks for observation of the pubertal development. Post-puberty heifers were in the luteal phase of their second cycle. There was no statistical difference in either BW (338 ± 54 and 363 ± 39 kg, $P = 0.38$) or CS (3.5 ± 0.4 and 3.8 ± 0.4, $P = 0.18$) between pre- and post-pubertal heifers.

Tissue samples (HYP, PIT, OVA, UTE, LIV, MUS and AT) were harvested as fast as possible after slaughter to preserve quality of RNA. In total, 96 tissue samples were available for RNA extraction (12 per tissue, except for OVA which had 24 samples available corresponding to the left and right ovaries). Total RNA was purified using a combination of RNeasy (QIAGEN, Australia) and TRIzol methods as previously described (Fortes *et al*. 2016; Nguyen *et al*. 2017a; Nguyen *et al*. 2018). All samples were passed quality control with RNA integrity numbers higher than 6.9.

The Illumina TruSeq sample preparation kit (Illumina, San Diego, CA) was utilized to construct cDNA libraries for each sample. Standard HiSeq 2000 sequencer analyser (Illumina, San Diego, CA) protocols were used to conduct RNA sequencing. Sequence reads were assembled and mapped to the annotated bovine genome (UMD3.1). Quality control and RNA-Seq expression analyses were performed using CLC Bio Genomic workbench software (CLC Bio, Aarhus, Denmark), with procedures described previously (Nguyen *et al*. 2017a; Nguyen *et al*. 2018). A threshold of the gene expression value (RPKM) ≥ 0.2 was utilized to annotated expressed genes (Mortazavi *et al*. 2008).

We applied "omics" pipeline developed by Nguyen *et al*. (2017b) to identify DE genes, TS genes, genes harbouring SNP associated with female fertility (heifer pregnancy, first service conception and age at first corpus luteum). From the predicted pre-pubertal and post-pubertal networks using PCIT which comprised DE, TS, TF and genes harbouring associated SNP (Reverter and Chan 2008), we applied an information lossless approach (Reverter and Fortes 2013) to explore the connectivity degree of all TF in the network. This approach allowed identification of the best trio of TF that, through their first neighbours, span most of the network topology. Finally, the list of DE genes (n = 2,116) was used as target list for functional enrichment analysis using Database for Annotation, Visualization, and Integrated Discovery (DAVID, Dennis 2013).

RESULTS AND DISCUSSION

An average of 60 million sequence reads were obtained for each individual sample. Previous studies demonstrated that approximately 30 million reads are sufficient to detect more than 90% of annotated genes in mammalian genomes (Lee *et al*. 2013; Wang *et al*. 2011). Despite the absence of a *Bos indicus* reference genome, our transcriptome data provided 60 to 70 % mapped reads. The relatively high number of sequence reads and mapped reads indicates that our data are adequate for differential expression studies.

A total of 2,116 DE genes, 624 TS genes, 186 TF and 179 genes harbouring SNP associated with heifer fertility traits were identified by comparing the pubertal status. Compared to a study by Cánovas *et al*. (2014) which used similar methods to identify genes in pre- and post-pubertal Brangus heifers, we found a higher number of DE genes, but lower numbers of TS genes, TF and genes harbouring associated SNP. The genetic makeup of Brangus heifers is 3/8 Brahman and 5/8 Angus. Differences in the breed type, the experimental design and sample size need to be considered when comparing the results of these two studies. Despite these discrepancies, comparing data from these two studies

could be useful to elucidate genes relevant for pubertal development in cattle, regardless of breed. Alternatively, specific genes delaying the pubertal process in Brahman heifers may be identified.

Based on gene ontology (GO) analysis of the 2,116 DE genes, we found enriched GO terms "G-protein coupled receptor protein signalling pathway", "regulation of hormone levels" and "steroid metabolic process". Metabolites and hormones are integrating peripheral signals for reproduction. Moreover, we also identified the most enriched biological process GO term: "immune response" (adjusted P = 8.3 x 10^{-13}). Reproduction is intimately connected to the immune function in women (Abrams and Miller 2011). The enrichment we found in cattle for the DE genes supports the idea of a relationship between reproduction and the immune system in cattle. The KEGG pathway neuroactive ligand–receptor interaction (adjusted P = 2.5 x 10^{-06}) has well known roles in puberty. This pathway comprises ligands and receptors noted to be involved in pubertal signalling such as glycoprotein hormones, alpha polypeptide, GABA receptor, OB-R, prolactin, prolactin receptor and growth hormone receptor (Ainu Husna *et al.* 2012).

The hub nodes of pre- and post-pubertal Brahman heifers sub-networks were *ZEB1*, *TEF* and *NFATC2* (Figure 1). Of note, *ZEB1* may control GnRH expression directly as well as indirectly (Messina *et al*. 2016), and was suggested as a candidate gene in a quantitative trait locus (QTL) study with pleiotropic effects on fatness, stature and reproduction in beef cattle (Bolormaa *et al*. 2014). Both our present study and the Brangus study (Cánovas *et al.* 2014) identified *ZEB1* as a key regulatory factor for bovine puberty. The gene *TEF* was reported as a transcription factor expressed in the pituitary gland during embryogenesis (Droplet *et al.* 1991). The initiation of *TEF* gene expression coincides with that of thyroid stimulating hormone beta (TSHβ). Droplet *et al*. (1991) reported that *TEF* can bind to and lead to effective transactivation of the TSHβ promoter. Thyroid hormones have a role in normal growth and reproductive function (Weber *et al*. 2013). The third TF of the best trio, *NFATC2*, belongs to the nuclear factor of activated T cells family that has been suggested to mediate GnRH action (Armstrong *et al.* 2009). These nuclear factors often generate signals in coordination with MAPKs (Macian 2005), which also play a role in GnRH regulation (Armstrong *et al*. 2009). In summary, our results amount to a growing body of evidence that supports these TF as important in the complex modulation of GnRH signaling and pubertal development.

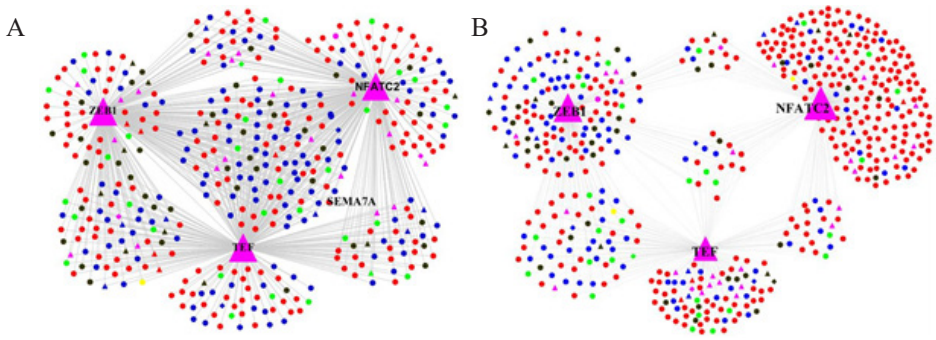

Figure 1. Sub-networks created with the best trio of transcription factors that span most of the network topology. A: pre-puberty network, B: post-puberty network. Genes are coloured according to their categories as follows: red = DE genes; pink = TF; blue = TS; dark brown = genes pertaining to two categories; and yellow = genes pertaining to three categories

Furthermore, examining the interaction between the best TF trio and other nodes in our sub-networks, we found that *SEMA7A* only interacted with the three TF in the pre-puberty network. In mice, during early development, loss of SEMA7A signaling can alter GnRH neuron migration and therefore lead to abnormal gonadal development and altered fertility (Messina *et al.* 2011). Protein and mRNA

expression of *SEMA7A* were observed in multiple neuronal systems (Pasterkamp *et al.* 2007). A study of the adult female rat brain suggested that *SEMA7A* was required for the neuroendocrine control of ovarian cycle (Parkash *et al*. 2015). Our result revealed only a slight and insignificant increase in the expression level of *SEMA7A* after puberty in HYP (FC = 0.2). However, significant DE *SEMA7A* ($P < 0.01$) was observed in the UTE (FC = -1.3) and PIT (FC = -0.9), representing a decrease in expression when progesterone signaling was present. We hypothesize that *SEMA7A* is regulated by the best trio of TF and could contribute to events leading to GnRH release in pre-pubertal Brahman heifers.

CONCLUSIONS

Our results provided potential candidate genes, pathways and networks related to pubertal development. Gene ontology terms and pathways identified from our target gene list might be informative to explain the molecular mechanisms involving in the onset of puberty in Brahman heifers. However, our current work was relying only on gene expression data and bioinformatics tools. Therefore, extensive functional experimental validation for these candidate genes is warranted.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by UQ to this research under a start-up grant awarded to Professor Stephen Moore.

REFERENCES

Abrams E.T and Miller E.M. (2011) *Am. J. Phys. Anthropol*. **146** Suppl 53: 134.

Ainu Husna M.S.S., Flatscher-Bader T., Lehnert S.A., Reverter A., Chan E.K., *et al*. (2012) *J. Trop. Agric. Food Sci*. **40**: 245.

Armstrong S.P., Caunt C.J., Fowkes R.C., Tsaneva-Atanasova K. and McArdle C.A. (2009) *J. Biol. Chem*. **284**: 35746.

Bolormaa S., Pryce J.E., Reverter A., Zhang Y., Barendse W., Kemper K., *et al*. (2014) *PLoS Genet*. **10**: e1004198

Cánovas A., Reverter A., DeAtley K.L., Ashley R.L., Colgrave M.L., *et al*. (2014) *PloS one*. **9**: e102551.

Dennis G.Jr., Sherman B.T., Hosack D.A., Yang J., Gao W., *et al*. (2003). *Genome Biol*. **4:** P3.

Drolet D.W., Scully K.M., Simmons D.M., Wegner M., Chu K.T., *et al*. (1991). *Genes Dev*. **5**: 1739.

Fortes M.R.S., Nguyen L.T., Weller M.M.D.C.A, Cánovas A, Islas-Trejo A., *et al*. (2016) *J. Anim. Sci*. **94**: 3693.

Lee H-J., Jang M., Kim H., Kwak W., Park W., Hwang J.Y., *et al*. (2013) *PloS one*. **8** :e66267.

Macian F. (2005) *Nat. Rev. Immunol*. **5**: 472.

Messina A., Langlet F., Chachlaki K., Roa J., Rasika S., *et al*. (2016) *Nat. Neurosci*. **19**: 835.

Messina A., Ferraris N., Wray S., Cagnoni G., Donohue D.E., *et al*. (2011) *Hum. Mol. Genet*. **20**: 4759.

Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B. (2008) *Nature methods*. **5**: 621.

Nguyen L.T., Reverter A., Cánovas A., Venus B.K., Islas-Trejo A., *et al*. (2017a) *J. Anim. Sci.* **95**: 599.

Nguyen L.T., Reverter A., Cánovas A., Venus B.K., Anderson S.T., *et al*. (2018) *Front. Genet.* **9** (87).

Nguyen L.T., Fortes M.R.S. and Reverter A. (2017b). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **22**: 289.

Parkash J., Messina A., Langlet F., Cimino I., Loyens A., Mazur D., *et al*., (2015) *Nat. Commun*. **6**: 6385.

Pasterkamp R.J., Kolk S.M., Hellemons A.J. and Kolodkin A.L. (2007) *BMC Dev. Biol*. **7**: 98.

Reverter A and Chan E.K. (2008) *Bioinformatics*. **24**: 2491.

Reverter A and Fortes M.R. (2013) *J. Anim. Sci*. **91**: 530.

Wang Y., Ghaffari N., Johnson C.D., Braga-Neto U.M., Wang H., Chen R., *et al*. (2011) *BMC Bioinformatics.* **12**: S5.

Weber G., Vigone M.C., Stroppa L., Chiumello G., (2003). *J. Pediatr. Endocrinol. Metab*. **16** (Suppl 2): 253.