Proc. Assoc. Advmt. Anim. Breed. Genet: 24: 357-360

INTEGRATION OF GWAS, NETWORK AND PATHWAY ANALYSIS REVEALS NOVEL INSIGHTS INTO THERMOTOLERANCE IN BEEF CATTLE

K.M. Sarlo Davila[1], E.E. Rodriguez[1], A.N. Nunez Andrade[1], G. Zayas[1], D. Rodriguez[1] and R.G. Mateescu[1]

[1] Department of Animal Sciences, University of Florida, Gainesville, FL, USA

SUMMARY

Thermotolerance, the ability to maintain production under heat stress conditions, is a complex trait determined by many component traits. Recent approaches combining traditional genome wide associations studies (GWAS) with gene network interactions theory could be more efficient in dissecting the genetic architecture of complex traits such as thermotolerance. Genes in common between several different gene ontology (GO) term groups might point towards key regulator genes with a greater impact on the thermotolerance complex. Highly connected genes identified in this analysis include *SYK, NOS2,* and *CD36.* While these genes have not been previously associated with thermotolerance, they have been associated with adaptation to other extreme environments including cold climates and high altitudes. These results indicate that there may be crucial genetic architecture responsible for environmental adaptation regardless of the nature of the challenging environment.

INTRODUCTION

Thermal stress in hot and humid conditions limits beef cattle production. Over 65% of the world's cattle (beef and dairy) reside in tropical or subtropical climates known for their hot and humid conditions. Thermotolerance, the ability to maintain production under heat stress conditions, is a complex trait determined by many component traits. Component traits related to heat loss are particularly of interest as there is a strong correlation between production level and metabolic heat production (Renaudeau *et al*. 2012). Greater capacity for heat loss rather than a lower metabolic heat level may allow cattle to maintain elevated production levels in the presence of heat stress (O'Brien *et al*. 2010). Many of the component traits that impact an animal's ability to lose heat are found at the hair-skin interface. Sweating capacity is of great importance as animals lose a majority of their heat through sweating when heat stress conditions become severe (Finch 1986). However, hair characteristics impact the effectiveness of sweating. Short, sleek hair coats allow for effective evaporative cooling during sweating, as well as reflect a greater proportion of solar radiation and facilitate conductive and convective heat flow (Hansen 2004). Recent approaches combining traditional GWAS with gene network interactions theory could be more efficient in dissecting the genetic architecture of complex traits such as thermotolerance. One advantage of association weight matrix/partial correlation information theory (AWM/PCIT) methodology is the ability to include SNP with relatively small effects that do not reach genome-wide statistical significance but are potentially linked to elements controlling the trait of interest. It is well recognized that many elements with minor effects are usually not able to reach significance at the genome level but will be uncovered through a gene network when multiple correlated traits are used in the analysis. GO term analysis of significant genes can be used to explore the functional mechanisms underlying thermotolerance.

MATERIALS AND METHODS

The University of Florida Institutional Care and Use Committee approved the research protocol used in this study (Approval no. 201203578). This study utilized 2,409 commercial Brangus heifers from the Seminole Tribe of Florida, Inc. Samples were collected from 12 groups of 200 animals: 4 groups over 4 consecutive weeks in each (August 15- September 12), 4 groups over 4 consecutive

weeks in 2017 (July 31 – August 28), and 4 groups over 4 consecutive weeks in 2018(July 26-August 23). Heifers within a year were from the same cohort and approximately the same age (about 2 years old).

The length and diameter of the undercoat (shorter coat closer to the body of the animal) and topcoat (longer coat that covers the undercoat) measured as described by Sarlo Davila *et al*. (2019). Coat score was measured for each heifer while in the chute and scored as 1 = very smooth, 2 = smooth, 3 = long, and 4 = woolly, as described by Hamblen *et al*. (2018). Sweating rate was measured using a calibrated, digital moisture sensor (Vapometer, Delphin Tech. Ltd, Kupio, Finland) that determines trans-epidermal water loss. The Vapometer uses a closed system approach, free of ambient airflow, to measure ambient relative humidity and temperature. The average body temperature of each heifer for each THI class from 0600 to 2000 hour was used in a random regression mixed model to estimate the reaction norm parameters for each individual: an intercept (RN intercept) and a slope (RN slope), as described in (Mateescu *et al*. 2020). The RN intercept describes the body temperature when animals are exposed to low heat stress (THI of 74 to 76), and the RN slope describes the change in body temperature in response to an increase of 5 THI units.

DNA was extracted from blood samples and genotyped with the Bovine GGP F250 array (Illumina Inc., San Diego, CA, United States). GWAS was performed as described by Sarlo Davila *et al*. 2020). The p-values and additive genetic values for each SNP were obtained for each phenotype and used to construct the association weight matrix (AWM) (Reverter and Fortes 2013). The AWM approach was used to synthesize the results from the GWAS. Topcoat length was chosen as the key phenotype to describe the complex of traits related to both thermotolerance and production. An initial set of 620 SNP with additive effects for topcoat length were selected based on their raw $P < 0.005$. To build the AWM, a vector of posterior mean estimates of the 620 SNP effects from topcoat was enhanced with the vectors of effects of all the other 7 phenotypes. This 620 x 8 matrix of posterior mean estimates of SNP effects was used as the input for PCIT to detect similar effects for any SNP across multiple phenotypes. All SNP pairs within the matrix were tested for association with at least one other SNP in order to establish network connections. SNP pairs without a significant partial correlation to at least one other SNP were removed from the dataset and discarded from subsequent network association analysis as they would appear isolated. SNP were then replaced with the gene the SNP were located in (within 2.5 kb), resulting in a network of 363 genes.

GO term enrichment and clustering were performed on all annotated genes from the AWM as well as the top genes from the GWAS for each trait. Functional grouping based on kappa score and visualization in a functionally grouped network was performed using the ClueGO plugin (Bindea *et al*. 2009) in Cytoscape. A kappa coefficient of 0.4 was used as a threshold value.

Table 1. GWAS results for the eight thermotolerance traits

Trait	n	Mean	Min	Max	h^2	Top Genes
Topcoat length (mm)	2163	14.95	3.84	33.53	0.36	*PRLR*
Undercoat length (mm)	2163	6.98	2.19	23.11	0.22	*PRLR, PCCA*
Topcoat diameter (mm)	2163	0.16	0.039	0.70	0.11	*MYC, RUNDC3A*
Undercoat diameter (mm)	2163	0.14	0.039	0.795	0.08	*RUND3CA, MYC*
Coat Score	2397	1.35	1	4	0.23	*PRLR, MMP19*
Sweating Rate	1319	60.68	2.00	218.0	0.11	*UBE2D2, CLIC3*
RN Intercept	2067	38.75	36.54	39.74	0.18	*ANKH, HELB, MGAT4B, PLA2G4*
RN Slope	2067	0.22	-0.001	0.442	0.10	*SLC22A10, RUNDC3A, LRRC49*

RESULTS AND DISCUSSION

A total of 363 annotated genes were found to be associated with at least one other gene and had significant direct and partial correlations. This correlation network generated a gene network consisting of 363 genes (nodes) and 22,928 relationships (edges). The top connected genes from the AWM were *RUNDC3A, TIGD7, OR4F73, YIPF1*and *PTPN21.* A functionally grouped annotation network was created from a list of 374 genes, combining the AWM results with the top GWAS results (Table 1). A network (Figure 1) was developed and visualized using the ClueGO plug-in for Cytoscape. 233 genes were associated with 103 biological function GO terms and pathways, forming 16 functional groups. The most representative group of terms was "anion transport" (18.52%), followed by "positive regulation of myeloid cell differentiation" (16.67%), and "organic acid transport" (11.11%). Higher connectivity between GO terms with similar molecular functions is to be expected, however, a high priority in terms of future research will be placed on genes in common between several different GO term groups as these might point towards key regulator genes with a greater impact on the thermotolerance complex. Highly connected genes include *SYK(spleen associated tyrosine kinase), NOS2 (Nitic Oxide Synthase 2),* and *CD36 (thrombospondin receptor).* *SYK* and *NOS2* were both connected to seven different functional groups while was *CD36* connected to six.

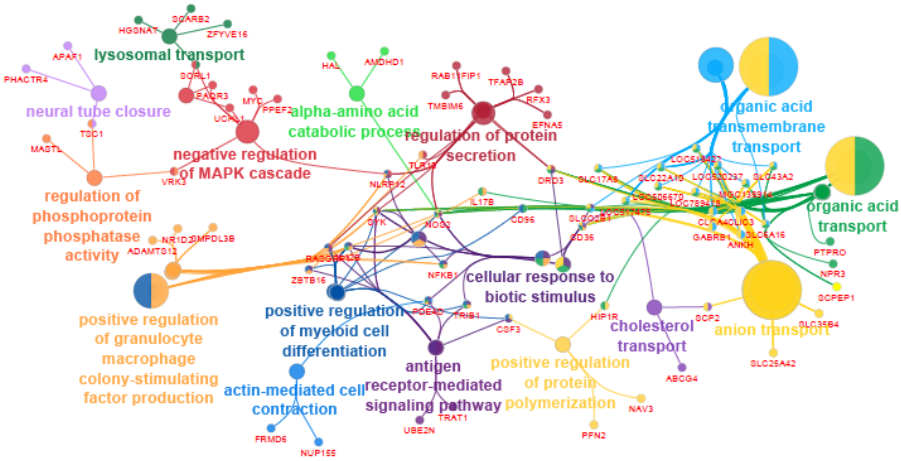

Figure 1. Functionally grouped network for thermotolerance. Biological process terms and genes (in red) as nodes. The node size represents the enrichment significance of the term

While none of these genes have been previously associated with thermotolerance, they have all been previously associated with adaptation to other harsh environments. *SYK* and *CD36* regulate brown adipose tissue and have been identified as candidate genes for adaptation to extreme cold. *SYK* was identified in a selection signature in two Russian breeds of cattle (Yurchenko *et al*. 2018) and *CD36* in Yanbian cattle (Shen *et al*. 2020). *SYK* is involved in brown adipocyte differentiation and *SYK* inhibition has been demonstrated to impair thermogenesis in mice (Knoll *et al*. 2017). *CD36* facilitates the uptake of energy substrates by brown adipose tissue and is essential for thermogenesis during cold exposure. *CD36* KO mice have been shown to have a drastically reduced body temperature after cold exposure (Putri *et al*. 2015). Brown adipose tissue is a key organ in non-shivering thermogenesis and helps cattle to conserve body heat in extremely cold environments. It is possible the absence of brown adipose tissue may help cattle lose excess body heat in hot environments.

NOS2 has been identified as a candidate gene for high altitude adaptation in cattle native to the

Ladakh region of India (Verma *et al*. 2018) as well as Zhangmu cattle native to China (Liu *et al*. 2020). *NOS2* upregulation was found to prevent hypoxia and is related to vasodilation. Enhanced expression of *NOS2* may increase the production of NO, resulting in vasodilation and increased blood flow to increase the O_2 supply (Verma *et al*. 2018). Increased blood flow to the skin also allows for effective heat dissipation via sweating (Finch 1986).

CONCLUSIONS

These results indicate that there may be crucial genetic architecture such as fat content and blood responsible for environmental adaptation regardless of the nature of the challenging environment, although the direction of selection for these traits changes with the environment. However, fat content, in particular, can also impact the production value of beef cattle as it affects meat quality. Further investigation of the impact of these traits on beef production is warranted.

REFERENCES

Bindea G., Mlecnik B., Hackl H., Charoentong P., Tosolini M., Kirilovsky A., Fridman W.H., Pagès F., Trajanoski, Z., and Galon, J. (2009) *Bioinformatics*. **25**(8): 1091.

Finch V.A. (1986) *J Anim Sci*. **62**(2):531.

Hamblen H., Hansen P.J., Zolini A.M., Oltenacu P.A., and Mateescu, R.G. (2018) *J Anim Sci*. **96**(8):3131.

Hansen P.J. (2004) *Anim Repro Sci*. **82:** 349.

Knoll M., Winther S., Natarajan A., Yang H., Jiang M., Thiru P., Shahsafaei A., Chavarria T.E., Lamming D.W., Sun L., Hansen J.B., and Lodish H.F. (2017). *Nature Comm*. **8**(1).

Liu X., Li Z., Yan Y., Li Y., Wu H., Pei J., Yan P. Yang R., Guo X., and Lan X. (2020) *Evol App*. **14**.

Mateescu R.G., Sarlo Davila K.M., Dikmen S., Rodriguez E., and Oltenacu P.A. (2020). *J Anim Sci*. **98**: 1.

O'Brien M.D., Rhoads R.P., Sanders S.R., Duff G.C. and Baumgard L.H. (2010) *Domestic Anim Endo* **38**: 86.

Putri M., Syamsunarno M.R.A.A., Iso T., Yamaguchi A., Hanaoka H., Sunaga H., Koitabashi N., Matsui H., Yamazaki C., Kameo S., Tsushima Y., Yokoyama T., Koyama H., Abumrad N. A., and Kurabayashi M. (2015) *Biochem Biophys Res Comm.* **457**: 520.

Renaudeau D., Collin A., Yahav S., de Basilio V., Gourdine J. L., and Collier, R. J. (2012) *Animal.* **6**(05): 07.

Reverter A. and Fortes M.R.S. (2013) *J Anim Sci*. **91**(2): 530.

Sarlo Davila K M, Howell A., Nunez A., Orelien A., Roe V., Rodriguez E., Dikmen S. and Mateescu R. G. (2020) *Anim Gen*. **51**

Sarlo Davila K.M., Hamblen H., Hansen P.J., Dikmen S., Oltenacu P.A. and Mateescu R. G. (2019) *J Anim Sci.* **97**(8): 3246.

Shen J., Hanif Q., Cao Y., Yu Y., Lei C., Zhang G. and Zhao Y. (2020) *Frontiers Gen.* **11**: 1.

Verma P., Sharma A., Sodhi M., Thakur K., Kataria R.S., Niranjan S.K., Bharti V.K., Kumar P., Giri A., Kalia S. and Mukesh M. (2018) *SciReports*. **8**(1): 1.

Yurchenko A., Daetwyler H., Yudin N., Schnabel R., Vander Jagt C., Soloshenko V., Lhasaranov B., Popov R., Taylor J. and Larkin D. (2018) *Sci Reports*. **8**(1): 1.