International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Identification and Prioritization of Genes Regulating Muscle Growth in Bovine Species

Dr. Mohan Mondal

Animal Physiology & Reproduction Laboratory, ICAR - National Dairy Research Institute, Eastern Regional Station, Kalyani, Nadia, West Bengal, India
Corresponding Author Email: *drmmondal[at]gmail.com*

Abstract: ***Introduction: Discovering novel genes regulating multigenic traits like meat quality in bovine species is challenging and required to perform urgently for further research in this field for development of novel technologies. Performing genetic studies frequently result in large lists of candidate genes of which only few can be followed up for further investigation. Gene prioritization establishes the ranking of candidate genes based on their relevance with respect to a biological process of interest, from which the most promising genes can be selected for further analysis. Objectives: a) To identify the genes that regulate multigenic traits like growth (meat quality) in bovine species, and b) To prioritize these identified genes involved in multigenic traits of bovine species. Results: In the present study, through using state - of - art bioinformatics tool, we have identified a total of 595 candidate genes specific to muscle growth in bovine species (cattle) for the first time. Furthermore, these identified genes of interest were also ranked based on their importance and priority. The 10 topmost ranked genes that influence body mass growth in bovine species are MSTN, DGAT1, GHR, PRNP, GH1, IGF1, SCD, VEGFA, LEP and CAPN1. Conclusions: As it is not possible to check each and every gene from a total of approximately 22, 000 genes of cattle for most important multigenic traits viz., muscle growth, prioritization of the candidate genes that have been established in the present study will be useful for further laboratory analysis of the limited ranked genes for better understanding of the body growth for attainment of early maturity for better meat production.***

Keywords: Muscle growth, Gene identification, Gene prioritization, Bovine, Multigenic traits

1. Introduction

Meat quality is one of the most important economic traits in farm animals particularly in mithun (Bos frontalis), which is exclusively reared for meat purpose. Meat quality trait has a multifactorial background and is controlled by an unknown number of quantitative trait loci (QTL). Genome research in farm animals progressed rapidly in recent years, moving from linkage maps to genome sequence. Phenotypic traits like muscle growth (meat quality) are generally complex traits and usually result from variation in many genes, each contributing only partially to the trait. Identification and molecular characterization of candidate genes underlying such traits remain a very challenging task because genetic mapping studies often have limited precision, leading to identification of relatively large genomic regions containing tens or even hundreds of genes.

High - throughput technologies have generated a vast amount of biological information. However, it remains a difficult task for biologists and clinical researchers to identify genes potentially important for a given biological process or disorders related to it based on these data. The researcher needs to determine which of those genes are of most interest and promising, and so the next step in the analysis is to prioritize the list and find the method to do so. To tackle this critical problem, the bioinformatics field has developed a number of solutions for gene prioritization [1]; these methods are typically based on the idea that genes whose expression patterns, subcellular localization, structural domains, molecular functions or physical interactions are similar to those known to be important for a given biological process or a pathology, are likely to play critical roles as well. Alternatively, genes can be prioritized on the basis of domain - specific knowledge for specific diseases and biological processes [2 - 6].

Discovering novel genes regulating multigenic traits like meat quality in bovine species is, therefore, challenging and required to perform urgently for further research in this field for development of novel technologies. Performing genetic studies frequently result in large lists of candidate genes of which only few can be followed up for further investigation. Gene prioritization establishes the ranking of candidate genes based on their relevance with respect to a biological process of interest, from which the most promising genes can be selected for further analysis. Therefore, the aim of the present study was to identify the genes that regulate multigenic traits like growth (meat quality) and prioritize these identified genes involved in multigenic traits of bovine species.

2. Methodology

Identification of genes that regulate multigenic traits like growth (meat quality) and different diseases in bovine species:

Data on the list of genes that control specific multigenic traits (phenotype) like growth (meat quality) in bovine species were collected from literature, gene and homology information from the MEDLINE, NCBI Gene, and HomoloGene databases, and also from Ge´nie web - based software [5]. The selected genes were ranked by their relevance to the query topic. The relevance is computed by a text mining algorithm using the number of abstracts related to the topic and the gene.

Prioritization of the identified genes involved in multigenic traits of bovine species:

The genes regulating muscle growth in bovine species (cattle) that were identified were used for prioritization through ranking. The Gene Prioritization Portal (http: //homes. esat. kuleuven. be/~bioiuser/gpp/tools. php) was

used [6] for prioritization of candidate genes for a specific type of multigenic phenotypic trait. Among the various tools available, we have mainly used two tools namely ENDEAVOUR and GPsy, which are the web - based software for prioritization of candidate genes specific to bovine species and available at http: //homes. esat. kuleuven. be/~bioiuser/endeavour/tool/endeavourweb. php (for many species: not exactly specific to bovine) and http: //gpsy. genouest. org (specific to bovine available), respectively [1].

3. Results and Discussion

Identification of candidate genes related to muscle growth in bovine species

For identification of candidate genes for the search 'skeletal muscle growth in cattle', the training set of 807 abstracts from PubMed were selected using the Ge´nie web - based software. The profile of the whole Medline database was considered as the background set. The test set was the protein - coding genes related to a taxonomic identifier (tax id = 9913). Cut offs were considered at $p<0.01$ for abstracts, and false discovery rate<0.01 for genes. A total of 595 genes were identified as shown below:

LEP, IGF1, DGAT1, SCD, GH1, PRNP, GHR, MSTN, CAPN1, INS, SREBF1, CAST, BOLA - DRB3, FASN, SLC11A1, PRL, TLR4, ADIPOQ, LEPR, IGF2, TNF, FABP4, IGF1R, PPARGC1A, IGF2R, IGFBP3, SPP1, BMP15, ABCG2, CXCL8, CSN2, STAT5A, POU1F1, TGFB1, MYOD1, LTF, CSN3, VEGFA, MC4R, CAPN3, NPY, GDF9 CSN1S1, IFNG, S1PR1, FSHR, FGF2, PPARG, HSPA1A, PAEP, UCP3, SIRT1, TG, PRG4, FABP3, GHRL, GNRHR, GHSR, NCAPG, ACAN, GHRH, FGF10, FGFR2, IFNT2, INSR, CARTPT, EDA, BMP4, POMC, PPARA, IGFBP2, TGFBR1, PGR, LHCGR, CSN1S2, TLR2, GART, ATP1A1, PC, SLC27A1, ZAR1, NR5A2, OLR1, CYP19A1, PTGS2, ACACA, KITLG, SLC5A1, CACNA2D1, DDX4, MBL1, SIRT2, RORC, IFNT3, CRH, DLK1, UTS2R, CEBPB, DGAT2, LALBA, PMEL, UCP2, TIMP3, SLC2A4, SERPINE2, NR1H3, FSHB, NR5A1, CASP3, CXCR1, MYF5, PRLR, CTNNB1, HSP90AB1, FBN1, MMP9, ELN, TMEM72, LOC789064, LOC101903326, MYH1, RCN1, TYRP1, VEGFB, WNT2, DIO3, CCAR2, HGD, MYBL2, CEACAM19, IFNGR2, STAT4, XKR4, SLC22A2, IGFALS, NFIX, ZBTB38, ME1, AKIRIN2, RTP3, FTO, PRDM16, LOC100196899, LOC100196900, LOC100295627, LOC100297748, C4A, PGF, DRD2, MBL2, CEBPA, GHRHR, STAT5B, COL10A1, IL12RB2, PROP1, SLC2A8, CD4, CL43, NLRP5, SLC7A7, MC3R, HSF1, IDO1, SLC37A2, AZGP1, TRMT61A, BCO2, ANXA9, SLC7A6, PLAG1, LHX3, CD180, LHX4, NPM1, SPRN, FAM167A, B3GNT5, AR, KIT, NMT1, CAPN2, ADRA1A, PRKAG1, ASIP, EGR1, PLTP, CRYZL1, SLC3A2, PRMT2, CPT1B, STAT1, SLC29A1, MYC, ITIH4, PRKAB1, SLC16A7, LOC100126230, HP, TSPY, LTBP2, BTN1A1, PGLYRP1, HGF, PAG2, F11, GRB14, SLC16A1, BOLA - NC1, ATP2A1, PGRMC2, FGF4, AMH, IL2, ACTA1, ADRB2, ADRB3, CCNA2, PSMC1, LOX, SRY, PLIN2, COMP, MC1R, SELL, ADAMTS4, FOXP3, SSLP1, MX1, STAR, HIF1A, HSPA1A, PECAM1, EGFR, PDGFB, ECE1, GPX1, GAPDH, SERPINE1, MMP14, SLC2A1, TLR1, ITGB2, LPL, CGN1, ISG15, BGN, MMP1, CYP17A1, LAP3, TIMP1, ESR1, ACTB, IL1B, ROCK1, FN1, NOS2, TIMP2, MMP2, ALB, ARSK, DNAJB2, NUDCD3, PICK1, DDX49, PAPD4, DHODH, DRAM1, NFYC, DDR1, VWA9, ZNF274, FKBP9, BIRC6, AIFM2, MTFR2, STX17, ARL15, SP2, RIPK2, EFNB1, NTM, FSTL1, SLC35A5, CCDC66, NME7, STIM1, CIAPIN1, C8H9orf91, ZEB1, MLST8, FAM73B, TRIM54, SCAMP1, PPME1, FAM76A, ELMOD3, LSM1, INPP4A, SLC3A1, SLC16A2, SMAD7, HOXC11, GPC6, PTPRR, TMEM229B, ADCK3, SLC27A6, ACSL1, TOP3B, PATL1, GPR180, LYAR, ARHGEF9, CPPED1, CCNE2, PTRF, AMOTL2, COL8A1, AGBL5, ABI3BP, HEYL, WNT16, C10H15orf41, NOG, SCAMP4, SIGMAR1, USP11, PRKAA2, BRINP1, FKBP2, TSPAN2, ZFP64, TRIM45, GEMIN4, KCTD11, LMAN2L, ZSCAN25, DCAF11, SLC7A1, PGBD1, PKLR, GLYR1, AP5M1, FANCF, ZFHX4, FGD5, PHF23, AICDA, ARHGAP24, LY6G6F, GJB3, BCORL1, GPR84, DNAJB5, FAM110A, KLHL22, NELFE, KANSL2, ZBTB6, PRMT6, SMAD4, HNRNPA0, ZNF45, CSRP3, MAPK1IP1L, SLC35C1, ARAF, LAX1, MINA, MAG, GPR61, MEAF6, DNAH1, CREM, RAB1B, TBCEL, CNOT4, DHX57, PPAPDC2, UBTD2, GDAP1L1, TMEM47, DUSP10, THSD1, NAPEPLD, SGCA, TNFSF8, ZFPL1, FEZF2, PLEKHO2, HMGB3, MED22, PTPN7, IDH3G, C8H9orf89, SYS1, MOB3C, IL23R, UBE2G2, PRSS16, MORN4, NXPH1, CTNNBIP1, JPH2, ASPHD1, FAM129A, CCDC130, KHK, ILF3, PPCDC, PUS10, BMP2, ACTR8, PEPD, EYA2, IRF5, HPS4, ADAT2, VASH1, CENPL, MGST2, FLOT2, EIF1AD, SDC2, VPREB3, FGF12, MRGPRF, FAM102A, GGT7, RNF220, GK5, TUT1, C28H1orf198, OPALIN, TTC38, RAD9B, TRAF3IP2, GOLGA7, CCDC8, NUDT10, COMMD4, H3F3B, LGI1, HSF2BP, SLC19A1, CXHXorf36, SCD5, TNFRSF12A, APOPT1, PSMF1, B4GAT1, IQCA1, ZNF3, SAMM50, NSUN4, PHLDA2, GATA6, WC - 7, CDC5L, MIS12, GPR89, LRRN2, FBXW2, IL10RB, UBE2L3, WDR60, ZFP2, C8H9orf64, ZFP36L2, GRN, GPR161, CARD9, KLHL12, DPH2, AQP3, C20H5orf28, ARHGEF39, HESX1, CPNE2, CXCR2, ERCC6L, SULT1C4, VPS13C, TSPAN18, CNTNAP3, AP2A1, PNPLA4, PJA1, VASN, TPK1, TAB3, HSD17B12, ERCC2, BTNL9, C29H11orf68, GTF3A, JMJD7, GPSM1, RAD54L, ARG2, MRPS9, CALM3, NAMPT, OGG1, CLPB, ACAA2, PLD2, FFAR2, PDK2, PLVAP, PINX1, RTCB, RCAN2, TNIP1, FLCN, IFIT2, SLC25A25, NR4A1, DDX19A, EPB41L5, CCR9, TRIM38, TAF6L, IFT57, POR, LAMC1, GPALPP1, KEAP1, DNAJC18, SWAP70, PDLIM7, EIF3E, MGAT1, RARA, ZNF410, FAM192A, ZBTB16, TMEM183A, FBLN5, MAPK13, ABCA1, RAPGEF3, ATL1, RBP1, ENPP6, DPAGT1, RALA, FBXO8, Sep - 04, ZCCHC17, RAB1A, TOLLIP, HSPB8, NFYA, CYP27B1, ASNSD1, BCS1L, FAM134B, MRPL35, MPZL2, MYOZ2, BMP7, PTEN, ANKRD49, KLHDC3, SLC22A5, MED17, NFIC, CDC42SE1, DVL2, NRXN3, PRSS35, MAPKAPK3, TCTA, TMBIM6, FEN1, SUPT4H1, SLMO2, GALM, FGF7, FAM213B, ERMN, RPH3AL, CEBPG, MMAB, PADI2, F13A1, AP5S1, CBX4, PLIN3, IGLL1, BOLA - N, LRWD1, AOC2

Prioritization of identified candidate genes that regulate multigenic traits like growth (meat quality) in bovine species

The genes regulating muscle growth in bovine species that were identified were used for prioritization through ranking using two web - based tools namely ENDEAVOUR and GPsy, which are available at http: //homes. esat. kuleuven. be/~bioiuser/endeavour/tool/endeavourweb. php and http: //gpsy. genouest. org, respectively [1].

A total of 595 candidate genes related to muscle growth in bovine species were identified using the Ge´nie web - based software. These genes were ranked (prioritized) based on their importance for the growth of muscle of bovine species (cattle) using two web - based tools namely Endeavour and GPSy. As Endeavour does not have the option for ruminants/bovine/cattle, we initially used GPSy, where option for bovine species is available. Finally, the prioritized candidate genes were also validated using Endeavour, which gives the similar output as that of GPSy. The top 50 prioritized genes are list as per their rank in the table 1.

Table 1: List of top 50 prioritized candidate genes regulating muscle growth in bovine species based on their rank

Rank	Gene ID	Symbol	PMIDs	Hits	FDR
1	281187	MSTN	33	23	1.07E - 56
2	282609	DGAT1	36	21	5.78E - 50
3	280805	GHR	25	18	3.70E - 46
4	281427	PRNP	72	22	1.97E - 44
5	280804	GH1	62	21	7.23E - 44
6	281239	IGF1	63	21	9.01E - 44
7	280924	SCD	32	18	1.53E - 43
8	281572	VEGFA	42	18	9.88E - 41
9	280836	LEP	59	19	6.69E - 40
10	281661	CAPN1	31	16	2.01E - 38
11	539361	SREBF1	15	13	1.61E - 36
12	281161	FGF2	35	15	1.06E - 34
13	281536	TLR4	35	13	1.86E - 29
14	281039	CAST	25	12	4.41E - 29
15	281152	FASN	12	10	3.41E - 28
16	282089	TGFB1	30	12	6.38E - 28
17	280943	TNF	48	13	1.83E - 27
18	280901	PRL	34	12	3.58E - 27
19	497205	LEPR	14	9	7.40E - 24
20	281740	CYP19A1	24	10	7.49E - 24
21	280867	PRG4	19	9	3.08E - 22
22	280828	CXCL8	21	9	9.32E - 22
23	282023	PTGS2	24	9	3.96E - 21
24	317698	IFNT2	27	9	1.35E - 20
25	407170	KDR	15	8	1.81E - 20
26	281948	NR5A1	8	7	2.42E - 20
27	281240	IGF2	29	9	2.57E - 20
28	280706	TG	18	8	1.10E - 19
29	281499	SPP1	34	9	1.25E - 19
30	281848	IGF1R	20	8	2.94E - 19
31	282375	STAT5A	10	7	3.04E - 19
32	281759	FABP4	11	7	8.08E - 19
33	281938	MYOD1	14	7	8.13E - 18
34	281993	PPARG	14	7	7.89E - 18
35	282261	IGFBP3	15	7	1.44E - 17
36	326285	FGF10	7	6	1.91E - 17
37	282574	GDF9	8	6	7.43E - 17
38	280794	FN1	19	7	1.03E - 16
39	280991	AKT1	19	7	1.01E - 16
40	280846	LTF	70	9	1.08E - 16
41	338446	PPARGC1A	9	6	2.01E - 16
42	281534	TLR2	22	7	3.15E - 16
43	281135	S1PR1	10	6	4.79E - 16
44	281914	MMP13	11	6	1.03E - 15
45	282315	POU1F1	11	6	1.01E - 15
46	287024	NOS3	109	9	5.93E - 15
47	282092	TIMP1	15	6	1.04E - 14
48	282530	BOLA - DRB3	39	7	2.46E - 14
49	282470	SLC11A1	17	6	2.46E - 14
50	327672	MAPK1	18	6	3.62E - 14

To the best our knowledge, this is the first study towards identification and prioritization of candidate genes regulating muscle growth in bovine species. Identifying causal genes that underlie complex traits such as muscle growth is a primary aim of genetic and biomedical studies. Genetic mapping of quantitative trait loci (QTL) and gene expression profiling based on high - throughput technologies are common first approaches toward identifying associations between genes and traits; however, it is often difficult to assess whether the biological function of a putative candidate gene is consistent with a particular phenotype [7 - 14]. Here, we have implemented a network - based muscle growth gene prioritization approach for ranking genes associated with quantitative traits in livestock species. The approach uses ortholog mapping and integrates information on muscle growth and disease or trait phenotypes, gene - associated phenotypes etc. [10 - 18]. It was used for ranking all known genes present in the cattle genome for their potential roles in bovine growth. The top ranked genes were highly enriched for pathways and biological processes underlying inflammation and immune responses, which supports the validity of our approach for identifying genes that are relevant to body growth of the animals. These gene - associated phenotypes were used for a local prioritization of candidate genes located in a QTL affecting the susceptibility to disease and influencing body mass growth. Our study provides a general framework for prioritizing genes associated with various complex traits in bovine species.

The results of our present study clearly showed that among the top 10 genes that influence body mass growth in bovine species are MSTN, DGAT1, GHR, PRNP, GH1, IGF1, SCD, VEGFA, LEP and CAPN1. Of which, importance of Growth hormone (GH1), Growth hormone receptor (GHR) Insulin - like Growth Factor - I (IGF1) and Leptin (LEP) genes for the body growth have been reported by several one - gene targeted studies. However, no efforts were made to identify and prioritize these genes based on their importance in a single study. Inhibition of Myostatin (MSTN) gene has been reported to enhance body growth in beef cattle [19 - 22] but its priority for growth studies was not known. Here, we described for the first time that MSTN gene holds top most priority for muscle growth in bovine species.

4. Conclusion

In the present study, through using state - of - art bioinformatics tool, we have identified the candidate genes specific to muscle growth in bovine species (cattle) for the first time. Furthermore, these identified genes of interest were also ranked based on their importance and priority. The 10 topmost ranked genes that influence body mass growth in bovine species are MSTN, DGAT1, GHR, PRNP, GH1, IGF1, SCD, VEGFA, LEP and CAPN1. As it is not possible to check each and every gene from a total of approximately

22, 000 genes of cattle for most important multigenic traits viz., muscle growth, prioritization of the candidate genes that have been established in the present study will be useful for further laboratory analysis of the limited ranked genes for better understanding of the body growth and also be helpful for conducting further advanced studies for development of technologies in the specific field, which is otherwise not possible with tens and hundreds of genes without gene prioritization, say for development of gene - based technology for attainment of early maturity for better meat production.

Declaration of Interest

Authors' have no conflict of interest confounding this research.

References

[1] Britto R, Sallou O, Collin O, Michaux G, Primig M and Chalmel F.2012. GPSy: a cross - species gene prioritization system for conserved biological processes—application in male gamete development. Nucleic Acids Research, 2012, 1–8. doi: 10.1093/nar/gks380

[2] Sun, J., Jia, P., Fanous, A. H., Webb, B. T., van den Oord, E. J., Chen, X., Bukszar, J., Kendler, K. S. and Zhao, Z. (2009) A multi - dimensional evidence - based candidate gene prioritization approach for complex diseases - schizophrenia as a case. Bioinformatics, 25, 2595–6602.

[3] Gajendran, V. K., Lin, J. R. and Fyhrie, D. P. (2007) An application of bioinformatics and text mining to the discovery of novel genes related to bone biology. Bone, 40, 1378–1388.

[4] Stein Aerts, Diether Lambrechts, Sunit Maity, Peter Van Loo, Bert Coessens, Frederik De Smet, Leon - Charles Tranchevent, Bart De Moor, Peter Marynen, Bassem Hassan, Peter Carmeliet & Yves Moreau.2006. Gene prioritization through genomic data fusion. Nature Biotechnology, 24 (5): 537 - 544.

[5] Jean - Fred Fontaine, Florian Priller, Adriano Barbosa - Silva and Miguel A. Andrade - Navarro.2011. Ge´nie: literature - based gene prioritization at multigenomic scale. Nucleic Acids Research, 2011, Vol.39, Web Server issue W455–W461. doi: 10.1093/nar/gkr246

[6] Tranchevent LC., Bonachela Capdevila F., Nitsch D., De Moor B., De Causmaecker P., Moreau Y. A guide to web tools to prioritize candidate genes. Brief. Bioinform. (2010)

[7] Tamayo, P. et al. Interpreting patterns of gene expression with self - organizing maps: methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. USA 96, 2907–2912 (1999).

[8] Stegmaier, K. et al. Gene expression - based high - throughput screening (GE - HTS) and application to leukemia differentiation. Nat. Genet.36, 257–263 (2004).

[9] Pixley, F. J. et al. BCL6 suppresses RhoA activity to alter macrophage morphology and motility. J. Cell Sci.118, 1873–1883 (2005).

[10] Galimi, F. et al. Hepatocyte growth factor is a regulator of monocyte - macrophage function. J. Immunol.166, 1241–1247 (2001).

[11] Brown, N. J. et al. Fas death receptor signaling represses monocyte numbers and macrophage activation in vivo. J. Immunol.173, 7584–7593 (2004).

[12] Scambler, P. J. The 22q11 deletion syndromes. Hum. Mol. Genet.9, 2421–2426 (2000).

[13] Baldini, A. Dissecting contiguous gene defects: TBX1. Curr. Opin. Genet. Dev.15, 279–284 (2005).

[14] Jerome, L. A. & Papaioannou, V. E. DiGeorge syndrome phenotype in mice mutant for the T - box gene, Tbx1. Nat. Genet.27, 286–291 (2001).

[15] Merscher, S. et al. TBX1 is responsible for cardiovascular defects in velo - cardio - facial/DiGeorge syndrome. Cell 104, 619–629 (2001).

[16] Lindsay, E. A. et al. Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. Nature 410, 97–101 (2001).

[17] Piotrowski, T. et al. The zebrafish van gogh mutation disrupts tbx1, which is involved in the DiGeorge deletion syndrome in humans. Development 130, 5043–5052 (2003).

[18] Rauch, A. et al. A novel 22q11.2 microdeletion in DiGeorge syndrome. Am. J. Hum. Genet.64, 659–666 (1999).

[19] Hayashi, T. et al. Identification and functional analysis of a caveolin - 3 mutation associated with familial hypertrophic cardiomyopathy. Biochem. Biophys. Res. Commun.313, 178–184 (2004).

[20] Zimprich, A. et al. Mutations in LRRK2 cause autosomal - dominant parkinsonism with pleomorphic pathology. Neuron 44, 601–607 (2004).

[21] Zuchner, S. et al. Mutations in the pleckstrin homology domain of dynamin 2 cause dominant intermediate Charcot - Marie - Tooth disease. Nat. Genet.37, 289–294 (2005).

[22] Munch, C. et al. Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. Neurology 63, 724–726 (2004).