Functional SNP in a polygenic disease induced by high-altitude in fattening Angus steers

Angela Cánovas

University of Guelph

67th Annual meeting EAAP, Belfast August 31st 2016

Single Genes > Genetical Genomics > System Biology

Markers

Linkage maps

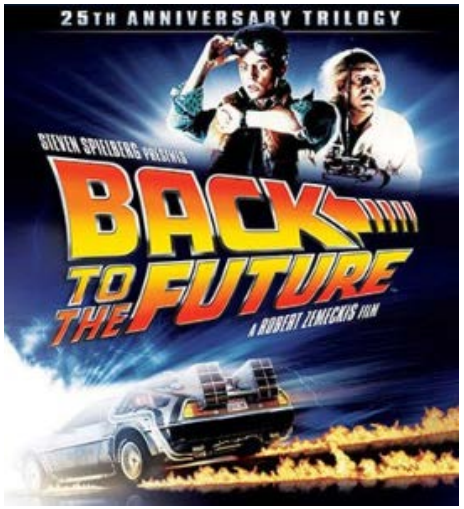
QTL

Expression Candidate Genes

qPCR

~ 2000 year

Single genes



Markers

Linkage maps

QTL

eQTL

qPCR

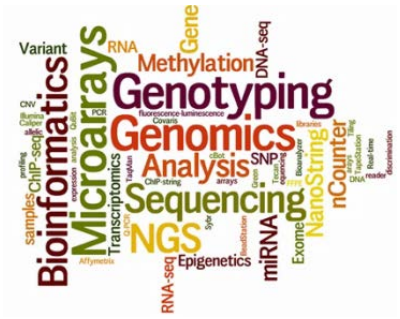
Global Expression

Microarrays

~ 2009 year

Genetical Genomics

Structural & Functional Genomic data



Present & Future Systems Biology

Statistical Methodologies

Structural Genomics

- QTL
- HD Markers
- Linkage maps
- GWAS
- CNV

Functional Genomics

- qPCR
- Global Expression
- Microarrays
- NGS
- RNA-Seq

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

How Does Genomics Accelerate Genetic Improvement

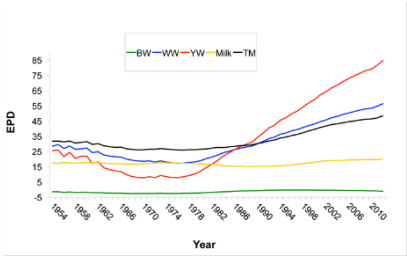
Why Genomics

Evolution of DNA Technology and Status of Ruminant Genome Sequence and Tools

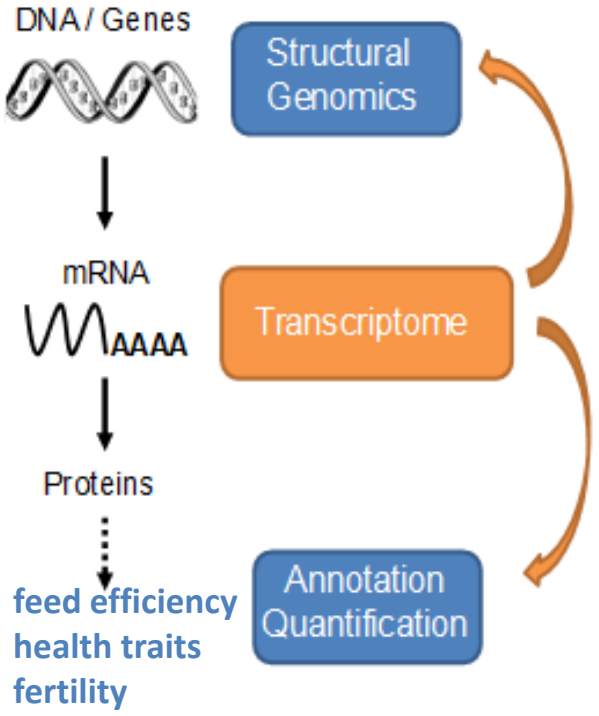
Item	Cow	Sheep	Goat	Yak
Reference genome	√	√	√	√
Additional animals sequenced	√	√		
50K SNP-chip	√	√	√	
HD SNP-Chip	√			
Other SNP-Chips	√			

?

Genetic Improvement



RNA-Sequencing



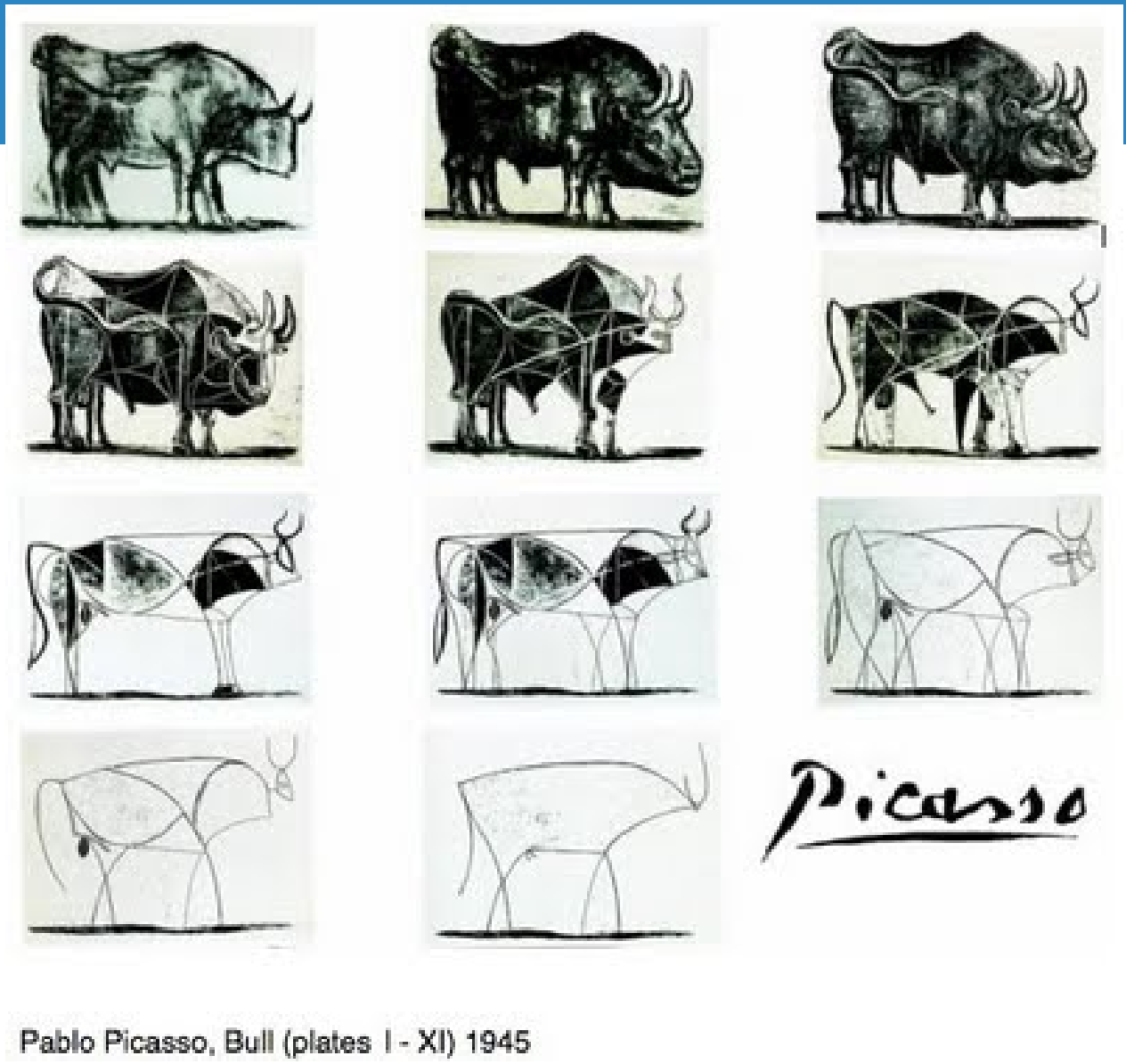
$$\Delta_{BV}/_t = \frac{r_{BV,\hat{BV}} \; i \; \sigma_{BV}}{L}$$

ΔBV x2+

↑ Accuracy of selection

↓ Reducing Generation Interval

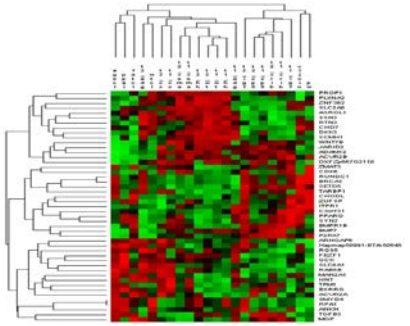
UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE



Pablo Picasso, Bull (plates I - XI) 1945

Analogy with Systems Biology where large amount of data could be integrate in order to extract the essential components that describe the biology affecting a complex phenotype

BMC Genomics

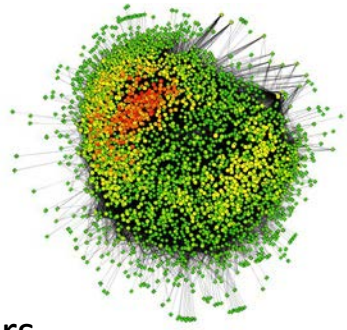


BMC Genomics. 2012 Jul 31;13:356.

Beyond differential expression: the quest for causal mutations and effector molecules.

Hudson NJ[1], Dalrymple BP, Reverter A.

PNAS
Proceedings of the National Academy of Sciences of the United States of America

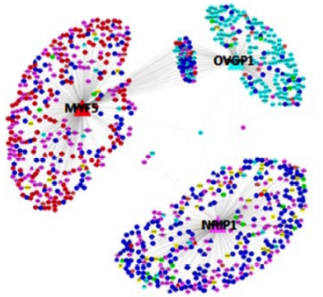


Proc Natl Acad Sci U S A. 2010 Aug 3;107(31):13642-7. doi: 10.1073/pnas.1002044107. Epub 2010 Jul 19.

Association weight matrix for the genetic dissection of puberty in beef cattle.

Fortes MR[1], Reverter A, Zhang Y, Collis E, Nagaraj SH, Jonsson NN, Prayaga KC, Barris W, Hawken RJ.

- Exploit the power of Multivariate analysis
- Combine the results from GWAS
- With the knowledge of Transcription Regulators
- In a Network Theory framework

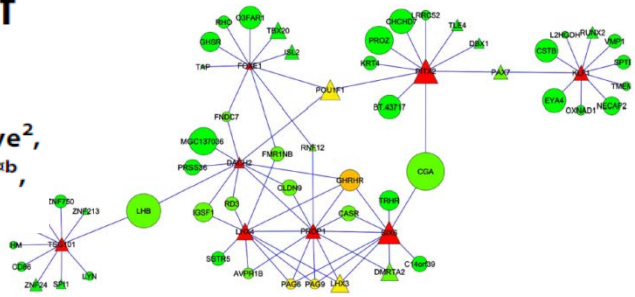


PLoS ONE 9(7): e102551.(2014)

PLOS | ONE

Multi-Tissue Omics Analyses Reveal Molecular Regulatory Networks for Puberty in Composite Beef Cattle

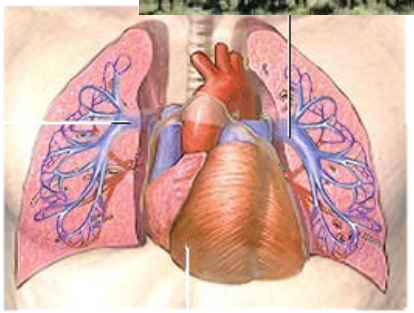
Angela Cánovas[1], Antonio Reverter[2], Kasey L. DeAtley[3¤a], Ryan L. Ashley[3], Michelle L. Colgrave[2], Marina R. S. Fortes[4], Alma Islas-Trejo[1], Sigrid Lehnert[2], Laercio Porto-Neto[2], Gonzalo Rincón[1¤b], Gail A. Silver[3], Warren M. Snelling[5], Juan F. Medrano[1], Milton G. Thomas[6*]



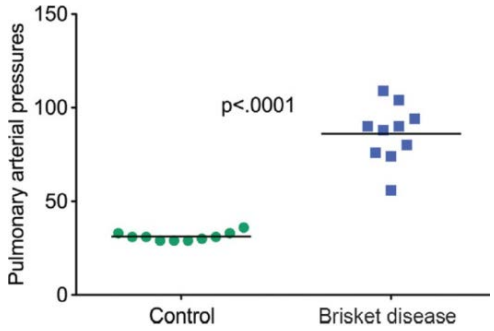
High-altitude or Brisket disease in beef cattle

- Economically important problem (>1,800 m).

- Associated with vascular inflammation of the pulmonary artery due to hypoxia.



pulmonary hypertension → congestive heart failure



- Heritable disease

Experimental design
RNA-Sequencing

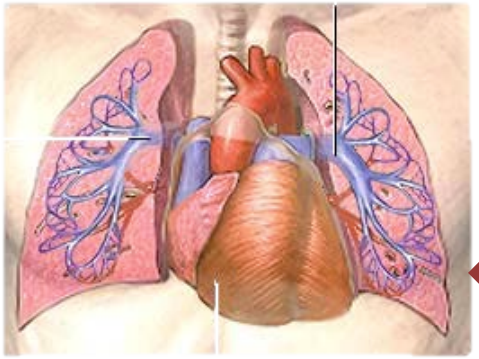
Experiment 1

Transcriptome from cultured bovine fibroblast of

Healthy calves

Control (n=6)

Normoxic conditions
1,500 m (5,000 ft)



Transcriptome from cultured bovine fibroblast of

Hypertensive calves

Hypertensive (n=6)

Hyporbaric hypoxia
4,500 m (15,000 ft)
14 days

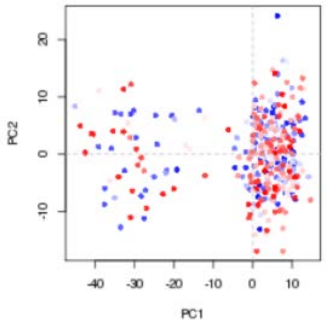
Experiment 2

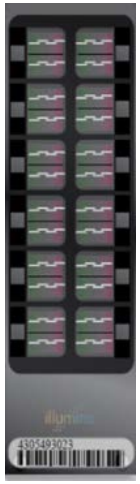
Transcriptome of

LPAP Angus cattle

LPAP (n=35)

Right Ventricle
Left Ventricle
Pulmonary Artery
Aorta
Muscle
Lung





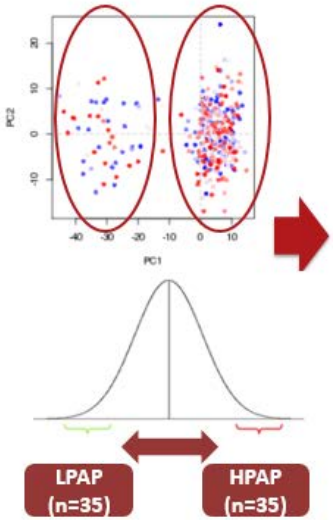
Transcriptome of

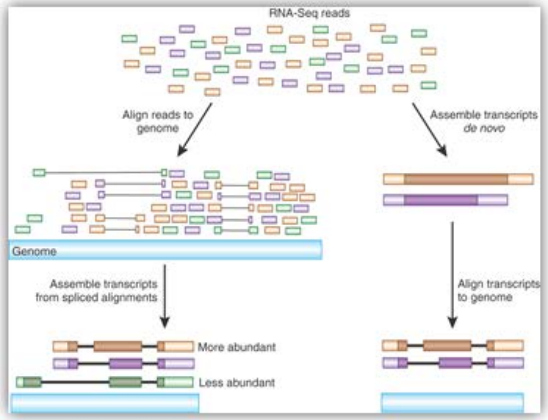
HPAP Angus cattle

HPAP (n=35)

Right Ventricle
Left Ventricle
Pulmonary Artery
Aorta
Muscle
Lung

RNA-Sequencing





Filtres & Quality Control Sequences

Mapping / Aligment
(paired –end 100 pb)

Bos taurus 3.1
(Release81)

Gene Expression
Structural Variation

"Genetical Genomics"

Integration of Structural & Functional Genomic data into Systems Biology

Splice Variants

Large gap mapping
Transcript discovery
RNA-Seq análisis

Statistical analysis
Differential Expression
LPAP vs **HPAP**

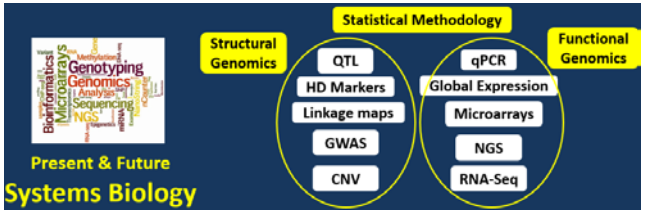
Metabolic Pathways

SNP Variants

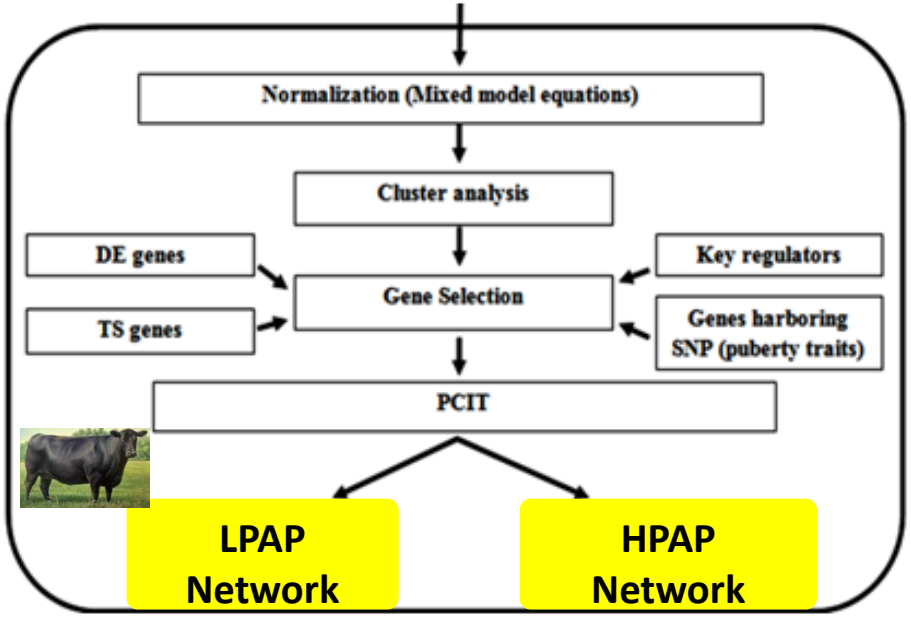
SNP group specific
(fixed/segreganting)

Variant Effect Predictor

Transcriptomics and Systems Biology of High Altitude Disease in Angus cattle

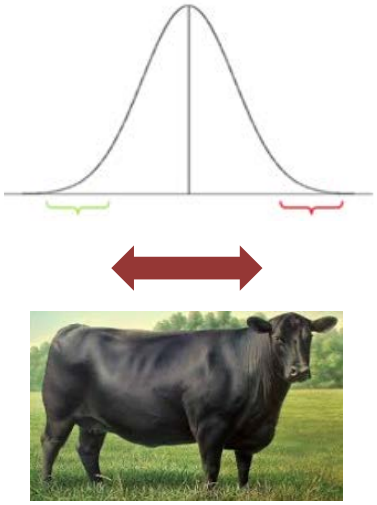


RNA-Sequencing
(70 samples)



- Differential Expression (RNA-Seq)
- Expr.Tissue specific (RNA-Seq)
- Splice variants Differentially expressed (RNA- Seq)
- Key Regulators - TF
- GWAS
- SNP Discovery

Differential Gene Expression (RNA-Seq)



RNA-Seq
DE analysis

HEALTHY vs **HYPERTENSIVE**
1,012 genes
(p<0.01; q<0.05; FC>2)

LPAP vs **HPAP**
(p<0.01; q<0.05; FC>2)

Cultured Fibroblast

Right Ventricle
Left Ventricle
Pulmonary Artery
Aorta
Muscle
Lung

LPAP vs HPAP

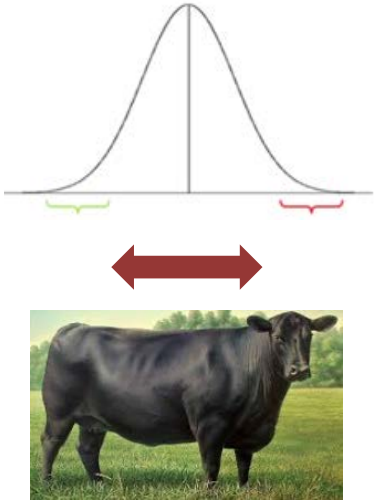
Tissue	DE gene (n)		
	p<0.05	p<0.01	p<0.01 + FC>2
Right ventricle	7,042	4,663	1,394
Left ventricle	791	158	20
Pulmonary artery	1,619	356	47
Aorta	5,685	2,590	1,173
Muscle	784	155	22
Lung	816	189	46

HPAP (Healthy) vs HPAP (Sick)

Tissue	DE gene (n)		
	p<0.05	p<0.01	p<0.01 + FC>2
Right ventricle	358	57	17
Left ventricle	559	107	25
Aorta	964	172	52
Muscle	474	99	40
Lung	372	84	26

Cánovas et al., XXVI Plant & Animal Genome. CA, 2016

Splice Variants Differential Expression (RNA-Seq)



HEALTHY vs **HYPERTENSIVE**
1,115 DE splice variants
(p<0.01; q<0.05; FC>2)

LPAP vs **HPAP**
(p<0.01; q<0.05; FC>2)

Cultured Fibroblast

Right Ventricle
Left Ventricle
Pulmonary Artery
Aorta
Muscle
Lung

Splice variants *DE*
LPAP vs **HPAP**

(p<0.01; q<0.05; FC>2)

ANNOTATED = 80% DE splice variants

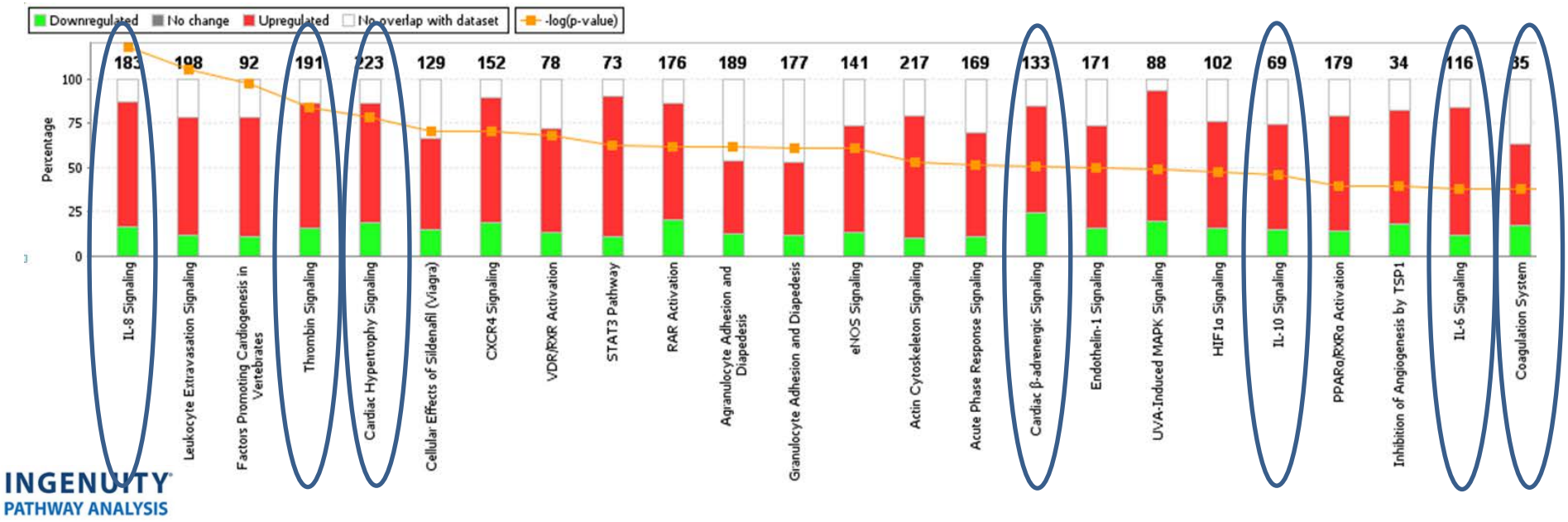
NOVEL = 20% DE splice variants

Blast2GO
IsoSeq PacBio

PACIFIC BIOSCIENCES®

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

Metabolic Pathway analysis

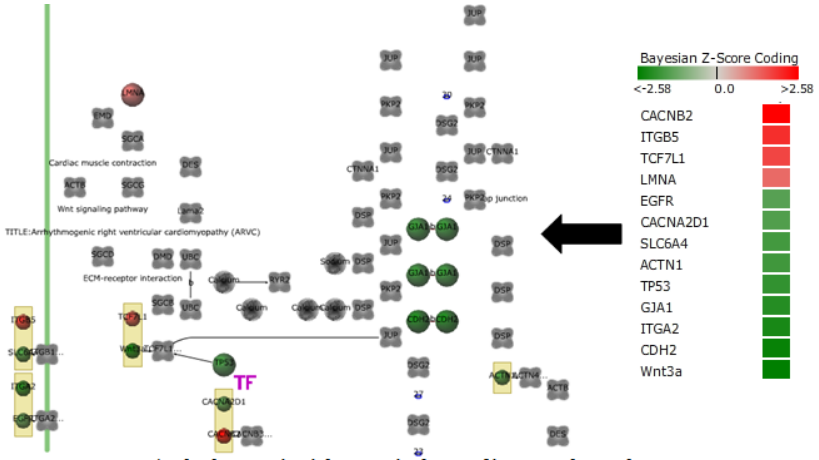


INGENUITY PATHWAY ANALYSIS

ASBAMC software

Bayesian approach

Adams et al., 2011



Arrhythmogenic right ventricular cardiomyopathy pathway.

Integrate metabolic pathway analysis to develop testable hypotheses in relation to the regulation of gene networks associated with brisket disease

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

Key Regulators – Transcription Factors

Right Ventricle
Left Ventricle
Pulmonary Artery
Aorta
Muscle
Lung

Regulatory Impact Factor metrics

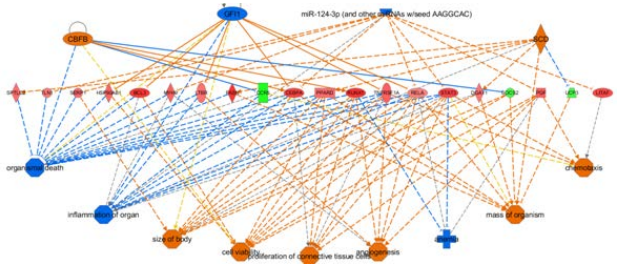
$$RIF1_r = \frac{1}{n_{DETS}} \sum_{j=1}^{j=n_{DETS}} x_j \times d_j \times DC_{rj}^2$$

$$RIF2_r = \frac{1}{n_{DETS}} \sum_{j=1}^{j=n_{DETS}} \left[\left(x_j^{PRE} \times r_{rj}^{PRE} \right)^2 - \left(x_j^{POST} \times r_{rj}^{POST} \right)^2 \right]$$

Reverter et al., 2008
Canovas et al., 2014

Right Ventricle

101 key regulators → regulating **705 genes differentially expressed** between LPAP and HPAP steers.



angiogenesis and atherosclerosis
cardiomyopathy (NFATC1)
movement of leukocytes and neutrophils (OLR1, PLAUR)
failure of heart (CTGF)
hypertrophy of heart ventricle (TREM1, GATA2)
vascularization (SYVN1)

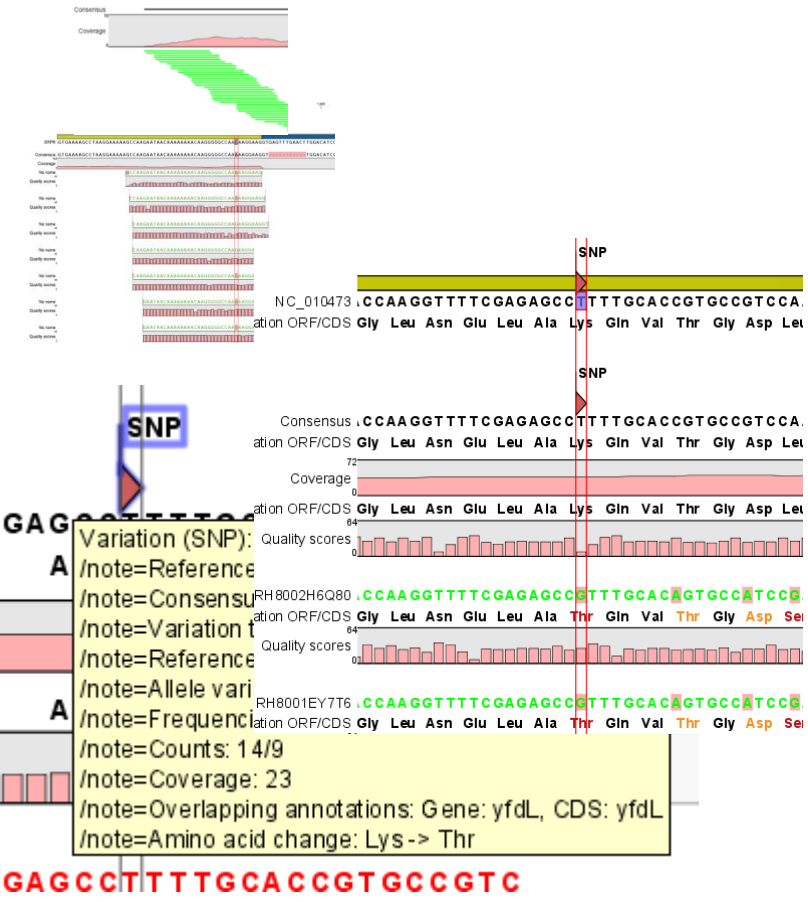
UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

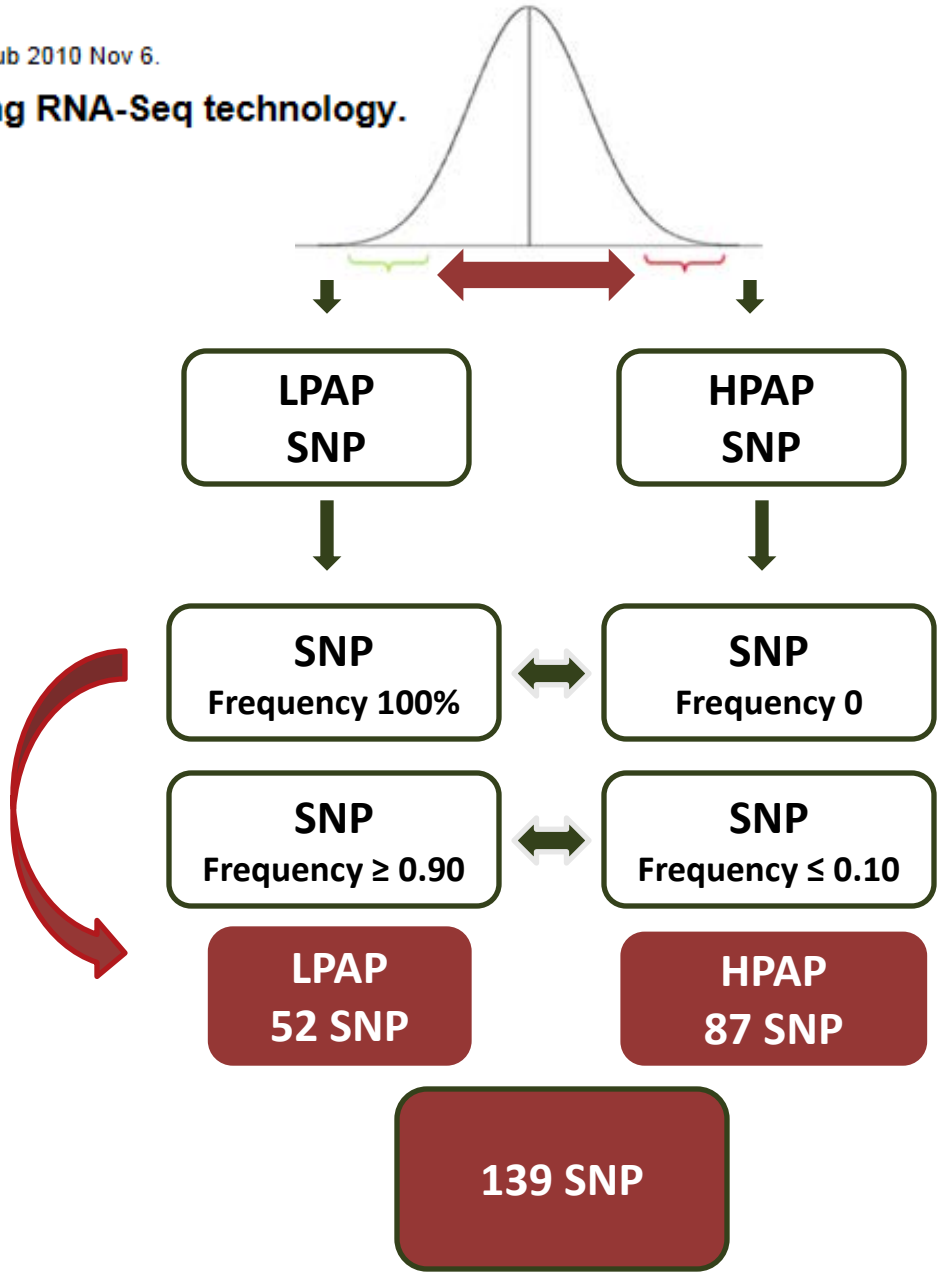
SNP Discovery

Mamm Genome. 2010 Dec;21(11-12):592-8. doi: 10.1007/s00335-010-9297-z. Epub 2010 Nov 6.

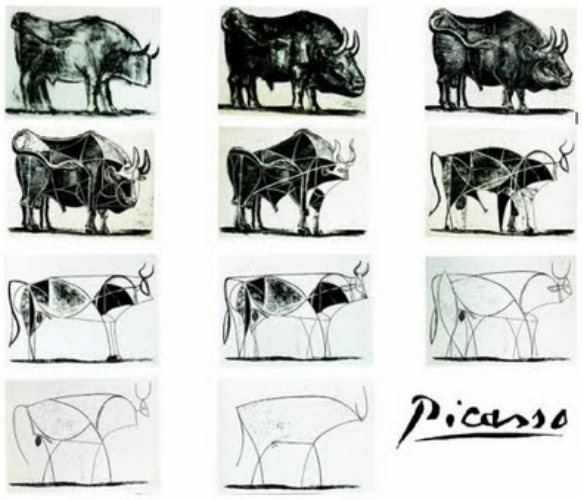
SNP discovery in the bovine milk transcriptome using RNA-Seq technology.

Cánovas A[1], Rincon G, Islas-Trejo A, Wickramasinghe S, Medrano JF.

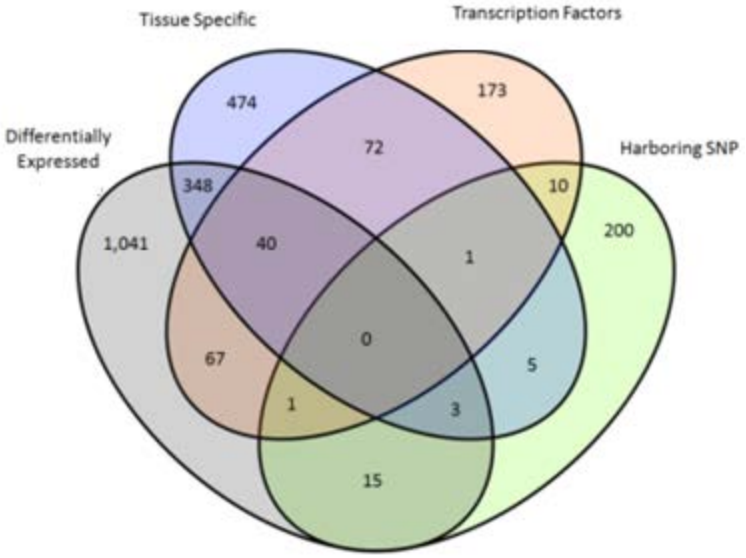




Transcriptomics and Systems Biology of High Altitude Disease in Angus cattle



Pablo Picasso, Bull (plates I - XI) 1945



- Differential Expression (RNA-Seq) ✓
- Expr.Tissue specific (RNA-Seq) ✓
- Splicing variants Differentially expressed (RNA- Seq) ✓
- Key Regulators – TF ✓
- SNP Discovery ✓
- Gene Networks ✓
- GWAS ✓

139 SNP Located in 36 Genes

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

Other works

nature genetics

A genetic mechanism for Tibetan high-altitude adaptation

Felipe R Lorenzo[1,19], Chad Huff[2,3,19], Mikko Myllymäki[4,19], Benjamin Olenchock[5], Sabina Swierczek[1], Tsewang Tashi[1], Victor Gordeuk[6], Tana Wuren[7], Ge Ri-Li[7], Donald A McClain[1], Tahsin M Khan[8], Parvaiz A Koul[9], Prasenjit Guchhait[10], Mohamed E Salama[11,12], Jinchuan Xing[2,13], Gregg L Semenza[14], Ella Liberzon[15,16], Andrew Wilson[17], Tatum S Simonson[2,18], Lynn B Jorde[2], William G Kaelin Jr[15,16], Peppi Koivunen[4,20] & Josef T Prchal[1,2,20]

PLOS | ONE

Genomic Scan Reveals Loci under Altitude Adaptation in Tibetan and Dahe Pigs

Kunzhe Dong[1☯], Na Yao[1☯], Yabin Pu[1], Xiaohong He[1], Qianjun Zhao[1], Yizhao Luan[2], Weijun Guan[1], Shaoqi Rao[2]*, Yuehui Ma[1]*

ARTICLE

Received 20 Jan 2015 | Accepted 6 Mar 2015 | Published 15 Apr 2015

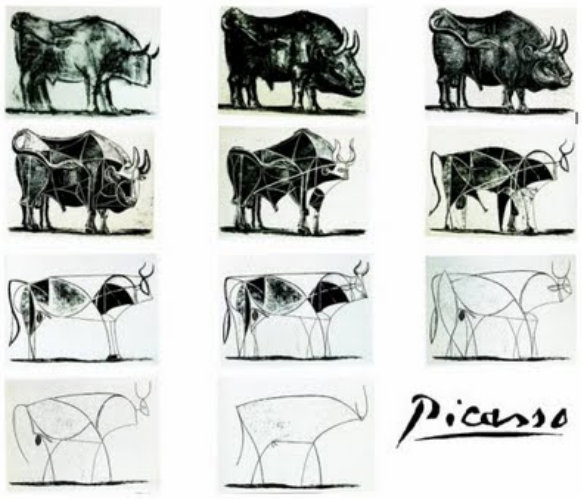
DOI: 10.1038/ncomms7863 OPEN

Increased prevalence of *EPAS1* variant in cattle with high-altitude pulmonary hypertension

John H. Newman[1], Timothy N. Holt[2], Joy D. Cogan[3], Bethany Womack[3], John A. PhillipsIII[3], Chun Li[4], Zachary Kendall[3], Kurt R. Stenmark[5], Milton G. Thomas[6], R. Dale Brown[5], Suzette R. Riddle[5], James D. West[1] & Rizwan Hamid[3]

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

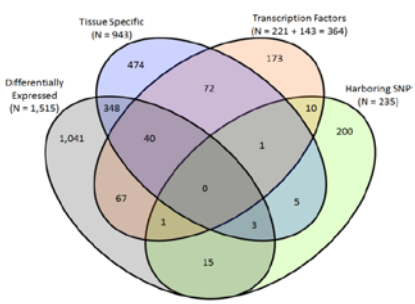
FUNCTONAL VALIDATION!!!!!

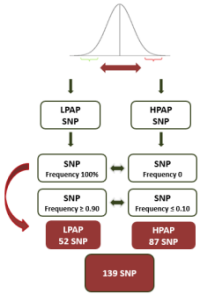


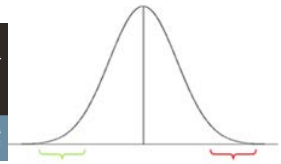
Pablo Picasso, Bull (plates I - XI) 1945

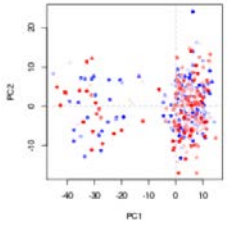
139 SNP
Located in 36 Genes

Functional Validation
Independent Population
1,000 steers approx.









… crossing fingers!!!!!

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

New Cattle Genome Assembly

*Slide courtesy of Juan F. Medrano

UMD3.1	Btau5.1	ARS-UCD v0.1
Reported April 2009 (Genome Biol)	Released 11/19/2015	In progress, update 7/10/2016
Based on 9x Sanger coverage	Based on 9x Sanger coverage initial	Based on ~80x PacBio data
WGS Dominette, BAC path Domino	+ 19x coverage P5 PacBio data	P6 chemistry, Falcon assembly
BAC end, RH map scaffolding and	BAC end, RH map, PBJelly2	(PacBio), Scaffolding: Dovetail HiRise,
human-cow synteny map.	scaffolding	Optical Map, Rec Map
75,618 contigs (97 kb contig N50)	**42,267 contigs (276 kb contig N50)**	**640 contigs (11.7 Mb N50)**
6337 scaffolds (6.4 Mb scaffold N50)	5,998 scaffolds (6.8 Mb scaffold N50)	29 scaffolds (110.9 Mb N50)

UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE

-OMICS technologies

- **New -OMICS Technologies to Understanding the Biological Processes and Network Pathways Associated with Cattle Growth and Health**

Phenotype

Genomics

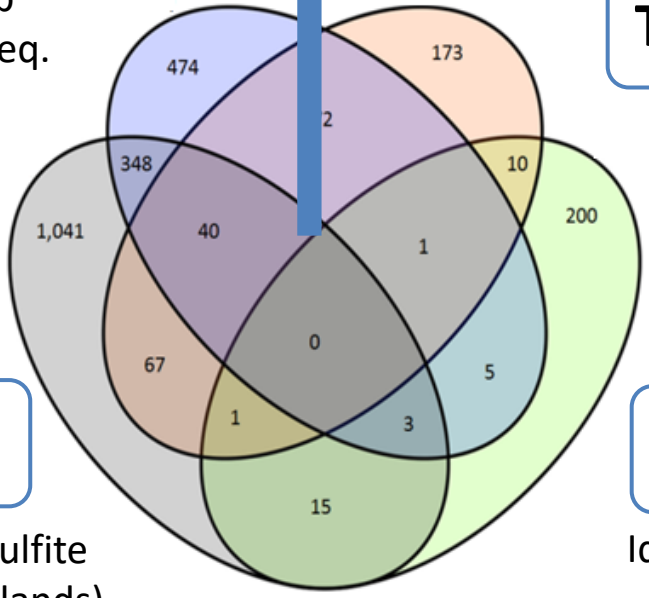
HD or LD SNP Chip
Whole Genome Seq.

Transcriptomics

Gene expression
Structural variation
-Indel
-Splice variants
-CNV

Metagenomics

Identification & Quantification
bacteria profile



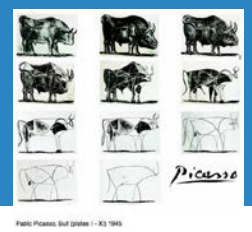
Methylomics

Genome Sequencing by bisulfite
Methylated pattern (CpG islands)

Metabolomics

Identification of biomarkers

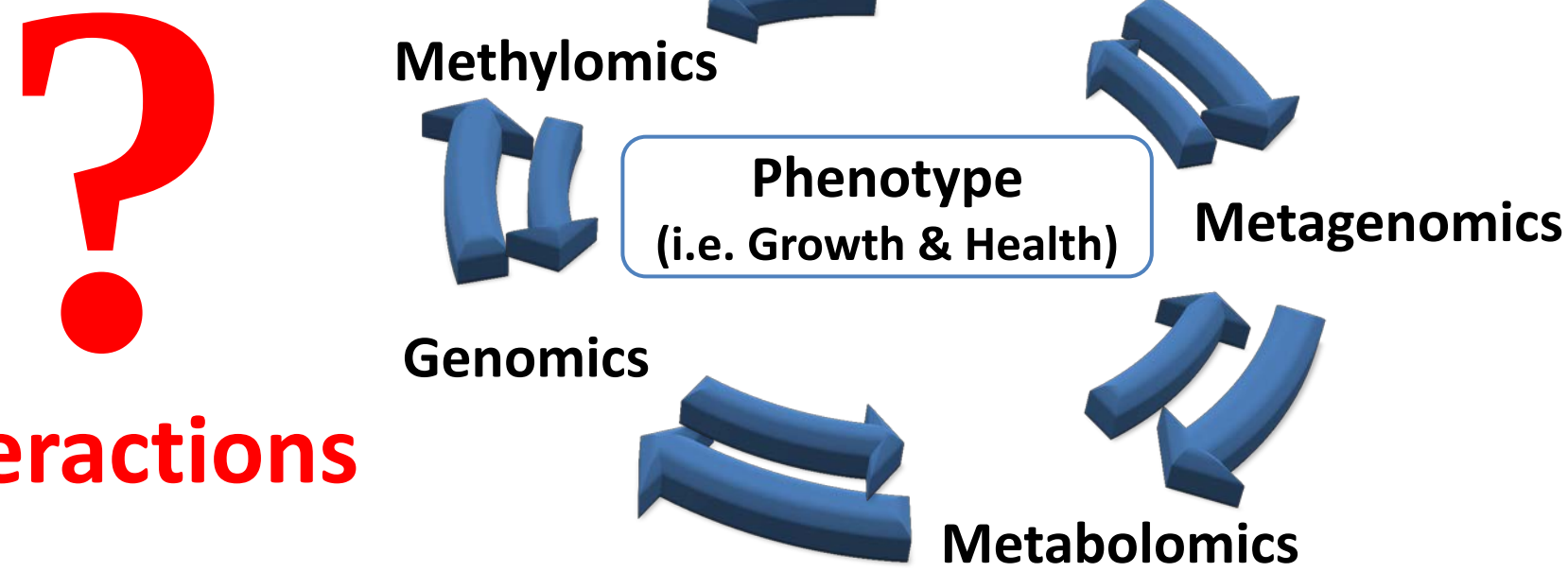
UNIVERSITY of GUELPH
CHANGING LIVES IMPROVING LIFE



-OMICS technologies

- **New -OMICS Technologies to Understanding the Biological Processes and Network Pathways Associated with Cattle Growth and Health**

Inferring causal phenotype networks using structural equation models



Acknowledgements

Milton G. Thomas
Joseph M. Neary
Tim Holt
Greta M. Krafsur
R. Mark Enns
Scott E. Speidel

Juan F. Medrano
Alma Islas-Trejo

Kurt R. Stenmark
Suzette K. Riddle
R. Dale Brown

Rebecca R. Cockrum

UC DAVIS
UNIVERSITY OF CALIFORNIA

Colorado State University

University of Guelph
Changing Lives Improving Life

University of Colorado Denver

Virginia Tech

Thanks!

Planking at **low altitude**
(sea level; San Diego, California)

Planking at **high altitude**
(~3,800 m; Rocky Mountain, CO)