Gota Morota January 2016

CONTACT Information A218f Animal Science Building Department of Animal Science University of Nebraska-Lincoln Lincoln, NE 68583-0908 USA E-mail: morota@unl.edu Phone: (402)-472-6031 WWW: morotalab.org

#### RESEARCH INTERESTS

I am a quantitative geneticist interested in incorporating statistics, machine learning, and bioinformatics to the study of animal genetics in the omics era. The core line of my research is connecting the quantitative genetics and animal breeding theories to currently available molecular information. I am particularly interested in statistical methods for prediction of complex traits using whole-genome molecular markers.

#### **EDUCATION**

### University of Wisconsin-Madison, Madison, Wisconsin USA

Ph.D., Animal Sciences, May 2014

- Dissertation: "Whole-genome Prediction of Complex Traits Using Kernel Methods."
- Advisor: Prof. Dr. Daniel Gianola
- Committee: Drs. Corinne D. Engelman, Guilherme J. M. Rosa, Grace Wahba and Kent A. Weigel

### University of Wisconsin-Madison, Madison, Wisconsin USA

M.S., Dairy Science, December 2011

- Thesis: "Application of Bayesian and Sparse Network Models for Assessing Linkage Disequilibrium in Animals and Plants."
- Advisor: Prof. Dr. Daniel Gianola
- Committee: Drs. Guilherme J. M. Rosa and Kent A. Weigel

# Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido Japan

B.S., Agricultural Science, March 2008

- Thesis: "Genetic Analysis of Threshold Traits."
- Advisor: Prof. Dr. Mitsuyoshi Suzuki

# Professional Positions

Department of Animal Science

University of Nebraska-Lincoln, Lincoln, Nebraska USA

Assistant Professor (70% Research & 30% Teaching)

08/2014 - Present

Work Experience Department of Animal Sciences

University of Wisconsin-Madison, Madison, Wisconsin USA

Graduate Research Assistant

06/2011 - 05/2014

Animal Genetics Research & Development Group

 ${\bf Zoetis,\ Inc.},$  Kalamazoo, Michigan USA

Quantitative Geneticist (student internship)

06/2013 - 11/2013

## Professional society memberships

- The International Society for Animal Genetics. 2015 Present
- The American Society of Animal Science. 2014 Present
- The Japanese Society of Animal Breeding and Genetics. 2013 Present
- The International Biometric Society (ENAR). 2012 Present

#### EDITORIAL ACTIVITIES

Reviewed for: Animal Genetics (1), Animal Production Science (1), BMC Genetics (2), Computers and Electronics in Agriculture (1), DNA Research (1), G3: Genes, Genomes, Genetics (1), Genetics (2), Genetics Selection Evolution (2), Journal of Animal Breeding and Genetics (2), Journal of Animal Science and Biotechnology (1), PLoS ONE (1), Poultry Science (1), Scientia Agricola (2)

Number of reviewed papers: 2012 (1), 2013 (1), 2014 (6), 2015 (10)

#### Preprints

15. <u>Morota G</u>, Beissinger TM, and Peñagaricano F. MeSH annotation of the chicken genome: MeSH-informed enrichment analysis and MeSH-guided semantic similarity among functional terms and gene products. bioRxiv. DOI: 10.1101/034975.

# PEER REVIEWED JOURNAL PAPERS

- 14. Abdollahi-Arpanahi R, <u>Morota G</u>, Valente BD, Kranis A, Rosa GJM, and Gianola D. 2016. Differential contribution of genomic regions to marked genetic variation and prediction of quantitative traits in broiler chickens. *Genetics Selection Evolution*. In press.
- 2015
- 13. Hu Y, Morota G, Rosa GJM, and Gianola D. 2015. Prediction of plant height in Arabidopsis thaliana from DNA methylation data. *Genetics.* 201:779-793. DOI: 10.1534/genetics.115.177204.
- 12. Valente BD, <u>Morota G</u>, Peñagaricano F, Gianola D, Weigel KA, and Rosa GJM. 2015. The causal meaning of genomic predictors and how it affects construction and comparison of genome-enabled selection models. *Genetics*. 200:483-494. DOI: 10.1534/genetics.114.169490.
- Morota G, Peñagaricano F, Petersen JL, Ciobanu DC, Tsuyuzaki K, and Nikaido I. 2015. An application of MeSH enrichment analysis in livestock. *Animal Genetics*. 46:381-387. DOI: 10.1111/age.12307.
- 10. Abdollahi-Arpanahi R, <u>Morota G</u>, Valente BD, Kranis A, Rosa GJM, and Gianola D. (2015) Assessment of bagging GBLUP for whole-genome prediction of broiler chicken traits. *Journal of Animal Breeding and Genetics.* 132:218-228. DOI: 10.1111/jbg.12131.
- Tsuyuzaki K, Morota G, Ishii M, Nakazato T, Miyazaki S, and Nikaido I. 2015. MeSH ORA framework: R/Bioconductor packages to support MeSH over-representation analysis. BMC Bioinformatics. 16:45. DOI: 10.1186/s12859-015-0453-z.
- 2014
- 8. <u>Morota G</u> and Gianola D. 2014. Kernel-based whole-genome prediction of complex traits: a review. *Frontiers in Genetics*. **5**:363. DOI: 10.3389/fgene.2014.00363.

- 7. <u>Morota G</u>, Boddhireddy P, Vukasinovic N, Gianola D, and DeNise S. 2014. Kernel-based variance components estimation and whole-genome prediction of pre-corrected phenotypes and progeny tests for dairy cow health traits. *Frontiers in Genetics*. **5**:56. DOI: 10.3389/fgene.2014.00056.
- Morota G, Abdollahi-Arpanahi R, Kranis A, and Gianola D. 2014. Genome-enabled prediction of broiler traits in chickens using genomic annotation. BMC Genomics. 15:109. DOI: 10.1186/1471-2164-15-109.
- Abdollahi-Arpanahi R, Pakdel A, Nejati-Javaremi A, Moradi-Shahrbabak M, Morota G, Valente BD, Kranis A, Rosa GJM, and Gianola D. 2014. Dissection of additive genetic variability for quantitative traits in chickens using SNP markers. *Journal of Animal Breeding and Genetics*. 131:183-193. DOI: 10.1111/jbg.12079.
- 4. Abdollahi-Arpanahi R, Nejati-Javaremi A, Pakdel A, Moradi-Shahrbabak M, <u>Morota G</u>, Valente BD, Kranis A, Rosa GJM, and Gianola D. 2014. Effect of allele frequencies, effect sizes and number of markers on prediction of quantitative traits in chickens. *Journal of Animal Breeding and Genetics.* 131:123-133. DOI: 10.1111/jbg.12075.
- 3. <u>Morota G</u>, Koyama M, Rosa GJM, Weigel KA, and Gianola D. 2013. Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. *Genetics Selection Evolution*. **45**:17. DOI: 10.1186/1297-9686-45-17.
- 2. <u>Morota G</u> and Gianola D. 2013. Evaluation of linkage disequilibrium in wheat with an L1 regularized sparse Markov network. *Theoretical and Applied Genetics*. **126**:1991-2002. DOI: 10.1007/s00122-013-2112-y.
- 1. <u>Morota G</u>, Valente BD, Rosa GJM, Weigel KA, and Gianola D. 2012. An assessment of linkage disequilibrium in Holstein cattle using a Bayesian network. *Journal of Animal Breeding and Genetics*. **129**:474-487. DOI: 10.1111/jbg.12002.

# Papers in Proceedings

2014

2011

2013

2012

- 3. Gianola D, <u>Morota G</u>, and Crossa J. 2014. Genome-enabled Prediction of Complex Traits with Kernel Methods: What Have We Learned? In: *Proceedings*, 10th World Congress of Genetics Applied to Livestock Production. August 17-22, Vancouver, BC, Canada. PDF
- 2. Valente BD, Morota G, Rosa GJM, Gianola D, and Weigel KA 2014. Causal meaning of genomic predictors: Implication on genome-enabled selection modeling. In: *Proceedings*, 10th World Congress of Genetics Applied to Livestock Production. August 17-22, Vancouver, BC, Canada. PDF

1. Bueno Filho JS\*, <u>Morota G</u>\*, Tran Q, Maenner MJ, Vera-Cala LM, Engelman CD, and Meyers KJ. 2011. Analysis of human mini-exome sequencing data from Genetic & Analysis Workshop 17 using a Bayesian hierarchical mixture model. In: *BMC Proceedings*. **5**(Suppl 9):S93. DOI: 10.1186/1753-6561-5-S9-S93. \*equal contribution

Invited Presentations	$\label{lem:matter} \mbox{My PDF presentation slides are available on $\bf SlideShare~(http://www.slideshare.net/chikudaisei/).}$
2016	6. MeSH annotation of the chicken genome. The Plant and Animal Genome XXIV Conference. Town and Country Hotel, San Diego, CA. January 9-13.
2015	5. Inferring the impact of population stratification on genomic heritability using a reparameterized genomic best linear unbiased prediction model. Statistics Seminars. Department of Statistics, University of Nebraska-Lincoln. September 23.
	4. Quantitative genetics in the functional genomics era. Animal Breeding & Genetics Seminars. Department of Animal Science, Iowa State University. March 3.
2014	3. Quantitative genetics in the functional genomics era. Special Seminar. PIC, Inc., Hendersonville, TN. November 12.
2013	2. Whole-genome prediction of complex traits using kernel methods. Department of Animal Science, University of Nebraska-Lincoln. December 19.
2011	<ol> <li>Obihiro GCOE Animal Global Health Seminars. Obihiro University of Agriculture and Veterinary Medicine. January 7.</li> </ol>
Contributed Presentations	
2015	<b>6</b> . Quantitative genetics in the functional genomics era. Special Seminar. The National Institute of Agrobiological Sciences, Tsukuba, Japan. November 12.
	<b>5</b> . Quantitative genetics in the functional genomics era. Special Seminar. Laboratory of Biometry and Bioinformatics, The University of Tokyo. November 6.
	4. The impact of population stratification on genomic heritability. NCERA-225 Meeting. North Dakota State University. October 22-23.
	3. An application of MeSH enrichment analysis in livestock. ADSA-ASAS Joint Annual Meeting. Rosen Shingle Creek, Orlando FL. July 12-16.

2. Prediction of complex quantitative traits using functional annotations and bootstrap aggregating. Special Seminar. National Livestock Breeding Center, Shirakawa, Japan. January 10.

1. Application of Bayesian and Sparse Network Models for Assessing Linkage Disequilibrium in Animals and Plants. 26th International Biometric Conference. Kobe International Conference Center, Kobe Japan. August 26-31. http://secretariat.ne.jp/ibc2012/30Aug.html#aug-30-14:00-

Contributed36. \*Second Oral Prize Winners.

Posters

• Population stratification contribution to genomic heritability. Probabilistic Modeling in Ge-2015 nomics. Cold Spring Harbor Laboratory, NY. October 14 - 17. • Estimating genomic heritability in the presence of population stratification. NGS Field 4th Meeting. Tsukuba International Congress Center, Tsukuba, Japan. July 1-3. 2013 • MeSHR: R/Bioconductor package for finding statistically overrepresented MeSH terms in a set of genes. Annual Bioconductor Conference BioC 2013. July 18-19, Seattle, WA. https://secure.bioconductor.org/BioC2013/posters.php#8. • Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. Annual Bioconductor Conference BioC 2013. July 18-19, Seattle, WA. https://secure.bioconductor.org/BioC2013/posters.php#7. Intramural SEMINARS 2015 • Animal Breeding & Genetics Seminars. Department of Animal Science, University of Nebraska-Lincoln. September 15. 2014 • Animal Breeding & Genetics Seminars. Department of Animal Science, University of Nebraska-Lincoln. September 18. Ph.D. Thesis Defense. Department of Animal Sciences, University of Wisconsin-Madison. May 12. • Dairy Science Graduate Seminars. Department of Dairy Science, University of Wisconsin-Madison. February 14. 2013 • Special Seminar. Zoetis, Inc., Kalamazoo, MI. August 8. 2012 • Animal Breeding & Genomics Seminars. Department of Animal Sciences, University of Wisconsin-Madison. April 10. 2011 • Master's Thesis Defense. Department of Dairy Science, University of Wisconsin-Madison. December 5. 2010 • Animal Breeding & Genomics Seminars. Department of Dairy Science, University of Wisconsin-Madison. November 30. • Animal Breeding & Genomics Seminars. Department of Dairy Science, University of Wisconsin-Madison. March 23.

• Animal Breeding & Genomics Seminars. Department of Dairy Science, University of Wisconsin-

# TEACHING University of Nebraska-Lincoln, Lincoln, Nebraska USA

Instructor ASCI 896 Statistical Genomics

Madison. November 25.

Guest Instructor

ASCI 432/832 Genome Analysis - April 16

Spring, 2015

# University of Wisconsin-Madison, Madison, Wisconsin USA

Teaching Assistant

ANSCI/DYSCI 363: Principles of Animal Breeding

Spring, 2011

Teaching Assistant

ANSCI/DYSCI 361: Introduction to Animal and Veterinary Genetics

Spring, 2011

#### RESEARCH SUPPORT

#### Internal Funding

PI: Gota Morota

• Layman Seed Award - \$9,910.00

UNL

June 1, 2015 - May 31, 2016

Proposal: Cracking the Blackbox of Whole-genome Prediction: Genome Partitioning of Predictive Ability

Role: Principal Investigator

#### SERVICE ACTIVITIES

#### Animal Breeding & Genetics Group

• Animal Breeding & Genetics Seminars organizer Department of Animal Science, University of Nebraska-Lincoln

Spring 2016

• Animal Breeding & Genetics Seminars organizer Department of Animal Science, University of Nebraska-Lincoln

Fall 2015

# OSS

# CONTRIBUTIONS

#### R packages

• dkDNA: http://cran.r-project.org/web/packages/dkDNA/index.html

# Bioconductor packages

- meshr
- MeSH.db
- MeSH.AOR.db
- MeSH.PCR.db
- MeSH.XXX.eg.db (84 packages)
  - MeSH.Aca.eg.db
  - $\bullet$  MeSH.Aga.PEST.eg.db
  - MeSH.Ame.eg.db
- MeSH.Aml.eg.db

- MeSH.Ana.eg.db
- MeSH.Ani.FGSC.eg.db
- MeSH.Ath.eg.db
- MeSH.Bfl.eg.db

- MeSH.Bsu.168.eg.db
- MeSH.Bsu.TUB10.eg.db
- MeSH.Bta.eg.db
- $\bullet$  MeSH.Cal.SC5314.eg.db
- MeSH.Cbr.eg.db
- MeSH.Cel.eg.db
- MeSH.Cfa.eg.db
- MeSH.Cin.eg.db
- MeSH.Cja.eg.db
- MeSH.Cpo.eg.db
- MeSH.Cre.eg.db
- MeSH.Dan.eg.db
- MeSH.Dda.3937.eg.db
- $\bullet$  MeSH.Ddi.AX4.eg.db
- MeSH.Der.eg.db
- MeSH.Dgr.eg.db
- MeSH.Dme.eg.db
- MeSH.Dmo.eg.db
- MeSH.Dpe.eg.db
- MeSH.Dre.eg.db
- MeSH.Dse.eg.db
- MeSH.Dsi.eg.db
- MeSH.Dvi.eg.db
- MeSH.Dya.eg.db
- MeSH.Eco.55989.eg.db
- MeSH.Eco.CFT073.eg.db
- MeSH.Eco.ED1a.eg.db
- MeSH.Eco.HS.eg.db
- MeSH.Eco.IAI1.eg.db
- MeSH.Eco.IAI39.eg.db
- MeSH.Eco.K12.DH10B.eg.db
- MeSH.Eco.K12.MG1655.eg.db
- $\bullet \ \operatorname{MeSH.Eco.O127.H6.E2348.69.eg.db}$
- $\bullet$  MeSH.Eco.O157.H7.EDL933.eg.db
- MeSH.Eco.O157.H7.Sakai.eg.db
- MeSH.Eco.S88.eg.db
- MeSH.Eco.UMN026.eg.db
- MeSH.Eqc.eg.db

# Github

• https://github.com/morota

- MeSH.Gga.eg.db
- MeSH.Gma.eg.db
- MeSH.Hsa.eg.db
- MeSH.Laf.eg.db
- $\bullet$  MeSH.Lma.eg.db
- MeSH.Mdo.eg.db
- MeSH.Mes.eg.db
- MeSH.Mga.eg.db
- MeSH.Miy.eg.db
- MeSH.Mml.eg.db
- MeSH.Mmu.eg.db
- MeSH.Mtr.eg.db
- MeSH.Nle.eg.db
- MeSH.Oan.eg.db
- MeSH.Ocu.eg.db
- MeSH.Oni.eg.db
- MeSH.Osa.eg.db
- MeSH.Pab.eg.db
- MeSH.Pae.PAO1.eg.db
- MeSH.Pfa.3D7.eg.db
- $\bullet$  MeSH.Pto.eg.db
- MeSH.Ptr.eg.db
- MeSH.Rno.eg.db
- MeSH.Sau.USA300TCH1516.eg.db
- MeSH.Sce.S288c.eg.db
- MeSH.Sco.A32.eg.db
- MeSH.Sil.eg.db
- MeSH.Spo.972h.eg.db
- MeSH.Spu.eg.db
- MeSH.Ssc.eg.db
- MeSH.Syn.eg.db
- MeSH.Tbr.9274.eg.db
- MeSH.Tgo.ME49.eg.db
- MeSH.Tgu.eg.db
- $\bullet~$  MeSH. Vvi.eg.db
- MeSH.Xla.eg.db
- MeSH.Xtr.eg.db
- MeSH.Zma.eg.db

Additional training

2014-2015

• Participant of the Research Development Fellows Program (RDFP)

2014

• Participant of Fall 2014 Adopting Research Based Instructional Strategies for Enhancing (ARISE) Professional Development Programs - Just in Time Teaching (JiTT)

PARTICIPATION IN MEETINGS, SYMPOSIUMS AND WORKSHOPS

2015

- The 29th International Mammalian Genome Conference. Yokohama Port Opening Memorial Hall, Yokohama, Japan. November 8-11.
- DNA Technology: Where we've been, where we are, and where we're headed. The US Meat Animal Research Center, Clay Center, NE. October 19.
- GO-FAANG Workshop. National Academy of Sciences Building, Washington, DC. October 7-8.

2014

- Sheep Genomics Workshop. University of Nebraska-Lincoln. November 13-14.
- NCERA-225 Meeting. Implementation and Strategies for National Beef Cattle Genetic Evaluation. Bozeman, MT. October 23-24.
- 19th Summer Institute in Statistical Genetics: "Module 23: Advanced Quantitative Genetics". University of Washington. July 23-25. Taught by Mike Goddard and Peter Visscher.
- 19th Summer Institute in Statistical Genetics: "Module 19: Statistical & Quantitative Genetics of Disease". University of Washington. July 21-23. Taught by John Witte and Naomi Wray.
- UC Davis Bioinformatics Training Program: "Using Galaxy for Analysis of High Throughput Sequence Data". University of California, Davis. June 16-20. Taught by the Bioinformatics Core.
- Short course: "Evolutionary Quantitative Genetics". University of Wisconsin-Madison. May 19-23. Taught by Bruce Walsh.

2013

• Short course: "Statistical methods for prediction of complex traits using whole-genome molecular markers". University of Wisconsin-Madison. May 27-31. Taught by Daniel Gianola and Gustavo de los Campos.

- Short course: "Introduction to genome-enabled selection & Inferring causal phenotype networks using structural equation models". Kyoto University. August 31. Taught by Guilherme J.M. Rosa.
- Short course: "Identifying Genes for Complex and Mendelian Traits Using Next Generation Sequence Data". 26th International Biometric Conference. August 26. Taught by Suzanne Leal.

• Short course: "Programming and computer algorithms with focus on genomic selection in animal breeding". University of Georgia. May 15 - June 1. Taught by Ignacy Misztal, Shogo Tsuruta, Ignacio Aguilar, Zulma Vitezica, and Andres Legarra.

2009

Symposium: "Statistical Genetics of Livestock for the Post-Genomic Era (SGLPGE)". University
of Wisconsin-Madison. May 4-6.

2006

• Short course: "Estimation of Variance Conponents in Animal Breeding". Obihiro University of Agriculture and Veterinary Medicine. November. Taught by Shogo Tsuruta.

#### Computer Skills

- Programming Languages: Python, Fortran
- Statistical/Numerical Computational Tools: R, Octave
- Content-description Languages: XML, XHTML, CSS, LATEX
- Operating Systems: Linux and Mac OS X

#### Miscellaneous

- Languages: English and Japanese
- Courses taken for credits at the University of Wisconsin-Madison
  - Spring 2012
    - Animal Sciences 875-004: Topics in Analysis of Quantitative Genomic Data (Daniel Gianola)
    - Dairy Science 875-005: Parallel Programming & High Performance Computing (Xiao-Lin Nick Wu)
  - Fall 2011
    - Dairy Science 875-005: Molecular Aspects of Animal Breeding (Hasan Khatib)
    - Statistics 840: Statistical Model Building and Learning (Grace Wahba)
  - Spring 2011
    - Mathematics 609: Mathematical Methods in Systems Biology (Gheorghe Craciun)
    - Statistics 610: Introduction to Statistical Inference (Chunming Zhang)
    - Statistics 992-001: Statistical Methods for QTL Mapping (Karl Broman)
  - Fall 2010
    - Statistics 609: Mathematical Statistics I (Chunming Zhang)
    - Statistics 701: Applied Time Series Analysis, Forecasting & Control I (Yazhen Wang)
    - Statistics 775: Introduction to Bayesian Decision & Control (Kam-Wah Tsui)
  - $\bullet$  Summer 2010
    - Population Health Sciences 904: Analytic Methods in Genetic Epidemiology (Corinne Engelman, Karl Broman, Bret Payseur, Kristin Meyers)
  - Spring 2010
    - Animal Sciences 875: Linear Models with Applications in Biology and Agriculture (Daniel Gianola)
    - Statistics 850: Theory & Application of Regression and Analysis of Variance II (Wei-Yin Loh)
  - Fall 2009

- Computer Science 576: Introduction to Bioinformatics (Colin Dewey)
- Dairy Science 875-006: Design & Analysis of Microarray Experiments in Agriculture (Guilherme J. M. Rosa)
- Dairy Science 875-011: Introduction to Bayesian Data Analysis with R (Xiao-Lin Nick Wu)
- Genetics 629: Evolutionary Genetics (John Doebley, Bret Larget, Bret Payseur)
- Statistics 849: Theory & Application of Regression and Analysis of Variance I (Sunduz Keles)

#### • Summer 2009

- Computer Science 367: Introduction to Data Structure

#### • Spring 2009

- Agronomy 771: Experimental Design (Mike Casler)
- Agronomy 772: Applications in ANOVA (Mike Casler)
- Mathematics 222: Calculus and Analytic Geometry
- Statistics 771: Statistical Computing (Michael Newton)

## • Fall 2008

- Statistics 424: Statistical Experimental Design for Engineers (Peter Z. G. Qian)
- Statistics 541: Introduction to Biostatistics (Ismor Fischer)
- Zoology 645: Modeling in Population Genetics & Evolution (Andrew Peters)

## • Summer 2008

- Computer Science 302: Introduction to Programming
- Mathematics 431: Introduction to the Theory of Probability

REFERENCES References and additional information available upon request.