



Differential Evolution between Monotocous and Polytoeous Species

Hyeonju Ahn^a, Kyu-Won Kim^{1,a}, Hyeon Jeong Kim², Seoae Cho², and Heebal Kim*

Department of Agricultural Biotechnology, Seoul National University, Seoul 151-742, Korea

ABSTRACT: One of the most important traits for both animal science and livestock production is the number of offspring for a species. This study was performed to identify differentially evolved genes and their distinct functions that influence the number of offspring at birth by comparative analysis of eight monotocous mammals and seven polytoeous mammals in a number of scopes: specific amino acid substitution with site-wise adaptive evolution, gene expansion and specific orthologous group. The mutually exclusive amino acid substitution among the 16 mammalian species identified five candidate genes. These genes were both directly and indirectly related to ovulation. Furthermore, in monotocous mammals, the *EPH* gene family was found to have undergone expansion. Previously, the *EPHA4* gene was found to positively affect litter size in pigs and supports the possibility of the *EPH* gene playing a role in determining the number of offspring per birth. The identified genes in this study offer a basis from which the differences between monotocous and polytoeous species can be studied. Furthermore, these genes may harbor some clues to the underlying mechanism, which determines litter size and may prove useful for livestock breeding strategies. (**Key Words:** Monotocous, Polytoeous, Differential Evolution)

INTRODUCTION

Mammalian species can be divided into two groups, monotocous and polytoeous, by the number of progeny per birth. The mechanism determining this reproductive trait in each species has not been completely identified. Many researchers have studied either genes or factors affecting litter size in single species instead of directly focusing on the differences between monotocous and polytoeous species. For example, in commercial pig breeds, Chinese Meishan and Large White, it was shown that the estrogen receptor (ER) locus is associated with increased litter size (Rothschild et al., 1996). Female pigs which are homozygous for estrogen receptor locus alleles produced more offsprings than pigs which had the undesirable alleles

even though they had the same genetic background. The prolactin receptor gene was also identified to be related to the total number born (TNB) and number born alive (NBA) by least squares method in five PIC lines (Vincent, 1997). Another study focused on the genes of monotocous species to prove that natural mutations in an ovary-derived factor, such as *FecX¹* gene in Inverdale sheep, can lead to increased ovulation rates and infertility phenotypes in a dosage-sensitive manner (Galloway et al., 2000). In addition, Retinol-binding protein 4 (*RBP4*), estrogen receptor, and prolactin receptor genes were demonstrated to have a connection with litter size and the number of piglets born alive in German pig lines (Drogemuller et al., 2001).

In order to identify important genes that underlie the differences between monotocous and polytoeous species, a novel approach focusing on the evolutionary genetic differences between the two groups of species was applied. In this study, we conducted functional comparative genomic analyses between monotocous and polytoeous mammals in a number of different scopes: specific amino acid substitution with site-wise adaptive evolution, gene expansion and specific orthologous group. Through these analyses, we identified potential factors which may cause

* Corresponding Author: Heebal Kim. Tel: +82-2-880-4803, Fax: +82-2-883-8812, E-mail: heebal@snu.ac.kr

¹ Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 151-742, Korea.

² C&K Genomics, Seoul National University Research Park, Seoul 151-919, Korea.

^a Both authors contributed equally to this work.

Submitted Nov. 5, 2013; Accepted Dec. 13, 2013; Revised Dec. 16, 2013

the differences in offspring number between monotocous and polytocous species. For these monotocous and polytocous specific genes, we conducted a gene expansion study. We used 16 mammalian species in this study divided into eight monotocous species, seven polytocous species and an outgroup: human, chimpanzee, orangutan, macaque, mouse, rat, dog, panda, cat, horse, cow, dolphin, pig, Tasmanian devil, opossum, and platypus. The genes identified from the analyses were examined in light of previous studies to determine their roles in monotocous and polytocous species.

MATERIALS AND METHODS

Preparation of orthologs sequence for evolutionary analysis

We collected protein and mRNA sequences of 16 species, and orthologous gene information including in-paralogs of current leaf species from the phylogenetic topology referencing 20,317 human genes from Ensembl (Hubbard et al., 2002) REST API (<http://www.ensembl.org>). Human, chimpanzee, orangutan, macaque, panda, horse, cow and dolphin were selected as monotocous species, and mouse, rat, dog, cat, pig and Tasmanian devil were selected as polytocous species with the platypus as an outgroup. The phylogenetic tree of the 16 species was supported by Timetree (Hedges et al., 2006). One-to-many and many-to-many orthologs were revised to one to one orthologs due to the constraints of the site-wise selective pressure test in which only single gene for each species are accepted. To do this, we selected a representative gene from in-paralogs of the species (Figure 1). Although all in-paralogs arising after a duplication event within a species are orthologous to other species, the most conserved gene among the in-paralogs is likely to retain functional conservation to its orthologs. Therefore, we designated this gene as the representative gene among the in-paralogs for its orthologs. We constructed a gene tree for each one-to-many and many-to-many orthologous gene set by neighbor-joining method using PHYLIP software package (Felsenstein, 1989) with option 'O', fixing platypus as an outgroup species. The gene which had the shortest distance from the branching point of the duplication event in the corresponding species was selected as the representative gene among the in-paralogs. This method was used because the shortest branch length implies minimum divergence among the in-paralogs from the common branching point. All in-paralogous genes by definition share a common branching point after the duplication event within a species, and the genes diverged from the point of duplication event as time goes on. Hence, the gene with the shortest branch length is the most conserved gene while the branch length infers evolutionary distance from the point of duplication event. Finally, we

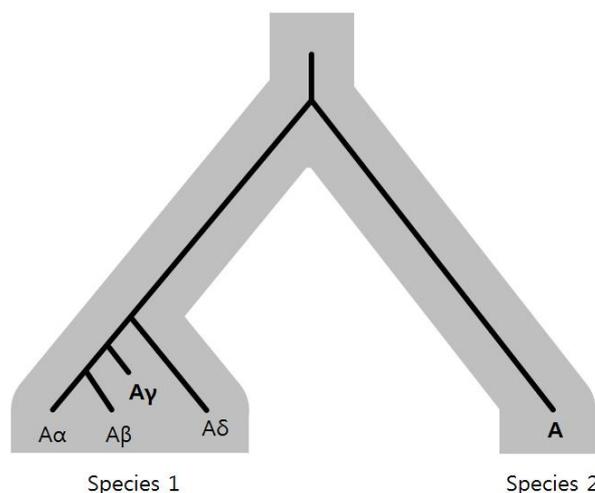


Figure 1. Schematic diagram for representative gene among in-paralogs in current exist species for comparative evolution analysis by dN/dS. $A\gamma$ is the representative gene having the minimum branch length from the most recent common duplication event shared by $A\alpha$, $A\beta$, $A\gamma$, and $A\delta$. The tree in gray is the species tree and the tree in black is the gene tree.

prepared 6,409 orthologous gene sets for the 16 species.

Parsimonious inference of convergent evolution

As monotocous and polytocous species are not monophyletic, the inference of episodic evolutionary history from which their traits formed through convergent evolution is somewhat complicated. We reconstructed the most likely history of adaptive evolution with a parsimonious approach. Problems with transition rates such as DNA transition rates can be solved by maximum likelihood estimation using substitution rate matrix, however, as episodic adaptive or purifying evolution has no transition rates, a parsimonious approach was used. Our pipeline generated all possible combinations of ancestral branches on the phylogenetic tree that have episodic adaptive evolutionary pressure that leads to either the acquisition or loss of the trait. Then we searched for the most parsimonious combinations that explained the present convergent evolution pattern for monotocous and polytocous species. Four branches were selected in which episodic adaptive evolutions occurred (Figure 2).

Mutually exclusive AA substitution for monotocous and polytocous species

Amino acid substitution sites in multiple sequence alignments, which are mutually exclusive between monotocous and polytocous species that could assert molecular convergent evolution for the trait, are termed 'mutually exclusive AA substitutions' in this study. Before testing for adaptive evolution, we filtered out orthologous

Gene symbol	Position in alignment	Mutually exclusive amino acid substitution type														Site-wise adaptive evolution					
		Human	Chimpanzee	Orangutan	Macaque	Mouse	Rat	Dog	Panda	Cat	Horse	Cow	Dolphin	Pig	Tasmanian devil	Opossum	Platypus	ω	$-2\Delta\ln(L_{H0}/L_{H1})$	FDR	BEB posterior probability
<i>TRMT11</i>	194	Y	Y	Y	Y	C	C	C	Y	C	Y	Y	Y	C	C	C	C	5.07	6.414	0.0142	0.999**
<i>CD22</i>	312	K	K	K	K	T	T	T	R	M	V	Q	R	L	T	T	T	6.27	7.961	7.97×10^{-3}	0.957*
<i>ARMCX3</i>	139	E	E	E	E	S	S	S	E	S	P	E	E	S	N	N	K				1.000**
	184	G	G	G	G	A	A	P	S	P	T	S	S	P	N	D	S	4.67	41.69	5.34×10^{-10}	0.977*
<i>LRRC19</i>	206	L	L	L	L	D	D	D	L	D	E	L	L	D	N	K	E				1.000**
	197	S	S	S	S	H	H	H	Y	R	C	C	C	R	N	N	N	23.3	13.84	4.97×10^{-4}	0.912
<i>TMEM214</i>	408	M	M	M	M	T	T	T	P	T	A	S	A	T	T	T	T	6.11	5.811	0.0159	0.801

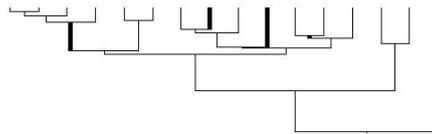


Figure 2. Genes showing site-wise adaptive evolution supported by mutually exclusive amino acid substitution from eight monotocous and seven polytocous mammals. Monotocous mammals are represented with a gray background color and bold letters. Phylogenetic relationship is presented below the table and the branches that have undergone adaptive evolution as implied by a parsimonious approach (see method) are represented with a thick line. ω : estimated dN/dS, H₀: null model, H₁: alternative model, BEB: Bayes empirical Bayes.

genes, which did not have these mutually exclusive AA substitutions. If there is a gene which is responsible for the number of offspring and had evolved differently between monotocous species and polytocous species, the gene would have converged within monotocous and polytocous species, and diverged between monotocous and polytocous species. In the evolution of a site, the aligned amino acid at the site do not need to be a single amino acid for converged evolution to have occurred because different amino acids can still lead to similar protein folding and hence similar gene function. However, the amino acid at the site needs to be different for divergent evolution since the same amino acid cannot lead to different functions when only the amino acid sequence is considered. So, we used the mutually exclusive AA substitutions as candidates of convergent evolution for our research purposes.

After aligning the orthologous gene sets of the 16 species with Prank (Löytynoja and Goldman, 2005) using the default options, the poorly aligned sites were filtered by Gblocks (Castresana, 2000). After filtering, 136 sets of orthologs remained.

Testing for site-wise adaptive evolution

We tested for site-wise adaptive evolution with the ratio of the rate of non-synonymous substitutions of codon sequence, which can give rise to functional change per non-synonymous site in the codon sequence to the rate of synonymous substitutions, which does not change the type of amino acid per synonymous site, denoted as dN/dS. In the case of neutral evolution, the dN/dS value converges to 1. For cases of adaptive evolution in which non-synonymous substitutions are promoted the ratio is over 1

while for cases of purifying selection in which non-synonymous substitutions are suppressed, the dN/dS value is below 1. The branch-site model was used for estimating ω , which is an estimated dN/dS value by maximum likelihood approach. In the model, there are three classes of ω values, $\omega_2 > 1$ for adaptive evolution, $\omega_1 = 1$ for neutral evolution and $\omega_0 < 1$ for purifying selection. We used “A model” of the branch-site model which estimates ω values under two hypotheses: H₀ and H₁. We fitted H₀ as a null hypothesis in which no adaptive evolution occurred for the branches from parsimonious inference of convergent evolution (see above method) fixing $\omega_2 = 1$ forcibly, and H₁ as an alternative hypothesis in which adaptive evolution occurred for the branches and $\omega_2 > 1$ is estimated. H₀ and H₁ were compared with log likelihood ratio test (LRT). After that, the p-values were adjusted by false discovery rate (FDR). Once the model had estimated ω_0 , ω_1 and ω_2 values, each amino acid site in the alignment was assigned into one of the three ω value by Bayes empirical Bayes (BEB) posterior probabilities.

We used PAML (Yang, 2007) for the implementation of “A model” of the branch-site model with options F3X4 for codon frequency for the 136 orthologs sets which have mutually exclusive AA substitutions. This resulted in 12 orthologs sets which were significantly estimated as having adaptive evolution. We selected the orthologous gene sets which have undergone adaptive evolution that were also on the site of mutually exclusive AA substitution to identify convergent evolution by adaptive selection pressure for the monotocous or polytocous traits. Through these steps, 5 orthologous gene sets were identified.

Gene expansion test

We investigated the monotocous specific and polytocous specific genes for expansion as it is possible that the trait was acquired by gene expansion. We collected 7,508 orthologous gene sets of the 15 species (this set included the platypus only as a reference and not as an outgroup so while the orthologous gene sets has genes of the platypus in it, a gene set did not have to include the platypus genes to be collected) including in-paralogs of current leaf species from the phylogenetic topology, in which at least one or more genes exist in a species from Ensembl REST. We counted the number of genes for each species per orthologous gene set. The number of genes between monotocous and polytocous species from each orthologous set was compared by a t-test to identify significant difference in the numbers.

Monotocous and polytocous specific gene sets

We collected monotocous and polytocous specific orthologous gene sets by selecting genes that only existed in the monotocous species or the polytocous species. We downloaded information of orthology relations including in-paralogs of the current leaf species from the phylogenetic topology for 319,682 genes of 16 species from Ensembl REST. We collected genes that have orthology relations for only monotocous or only polytocous species.

Since genes were clustered into orthologous gene sets, there were genes with the same information for its orthology relations; these were eliminated to reduce redundant information. After filtering, 10 monotocous specific orthologous gene sets and 12 polytocous specific orthologous gene sets were collected. We validated these gene sets by searching them against their counterpart's reference gene database. Each of the longest translation assumed to have the most informative translation among isoforms of the 10 monotocous specific orthologous gene sets was queried in the polytocous reference gene databases of NCBI (Pruitt et al., 2007) (<http://www.ncbi.nlm.nih.gov/>) through blastp query submission. The same was carried out for the 12 of polytocous specific orthologous gene sets. Considering that Ensembl database is one of the highest accessed database, it is likely that false positives are suppressed, however, for our purpose is to find the complement set against well-defined orthologous sets, false negatives should be avoided. False positives and false negatives are competitive concepts of each other. Therefore, we added a validation process on the specific orthologous gene sets.

RESULTS AND DISCUSSION

We explored genes which have evolved differently

between monotocous and polytocous species in a number of different scopes, specific amino acid substitution with site-wise adaptive evolution, gene expansion and specific orthologous group.

Site-wise adaptive evolution of monotocous species

Among the 6,409 orthologous gene sets of the 16 species, the genes that have an adapted site specific to monotocous species and are supported by a mutually exclusive monotocous amino acid substitution are shown in Figure 2. One of the identified genes is *TRMT11*, which shows a clear convergence between monotocous and polytocous species with a biallelic site. Monotocous and polytocous species each had a different amino acid even though the species were not monophyletic within their group. Furthermore, the gene *ARMXC3* and *LRRC19*, also showed convergence between the two groups with a small number of amino acids. Other genes also had mutually exclusive amino acid substitutions sites suggesting that these genes evolved differently between monotocous and polytocous species (see method).

For all five of the identified gene, the primates (human, chimpanzee, orangutan, and macaque) had an accelerated site with the same amino acid type (Figure 2). The shared amino acid type in the primates corresponds to the monophyly of the group. Mouse, rat and dog had an accelerated site within the five genes and shared the same amino acid type. The results of the parsimonious inference showed that the episodic adaptive evolution event supports the ancestral branches of the monotocous group. However, interestingly, the mouse and rat which are in the order Rodentia, and dog which is in the order Carnivora, are phylogenetically distance and yet showed strong convergence.

We found direct and indirect connections between the five identified genes and ovulation. *TRMT11* is a tRNA of methyltransferase, and interacts with E2 metabolites, which compounds NGTX (nongenotoxic carcinogens) as a catechol-O-methyltransferase (Jennen et al., 2010). NGTX contains tetra-chlorodibenzo-p-dioxin (TCDD) to induce interferon-inducible genes and genes associated with collagen synthesis or degradation in human amniotic epithelial cells (Abe et al., 2006). *CD22* is closely related to leukocyte and B lymphocyte and the immune cells are assumed to be strongly associated with the ovulation processes. *CD22* was expressed in B-lineage acute leukemia (Toba et al., 2003) and as a B lymphocyte adhesion molecule, *CD22* was found to interact with leukocyte common antigen CD45RO on T cells (Stamenkovic et al., 1991). In rabbits, T lymphocytes dramatically increased in the uterus after ovulation, and both before and after ovulation, the cells were observed frequently in mucosa of

Table 1. Gene families with the most significant difference in the number of expansion genes for monotocous and polytocous mammals¹

Type	Gene family		Human	Chimpanzee	Orangutan	Macaque	Mouse	Rat	Dog	Panda	Cat	Horse	Cow	Dolphin	Pig	Tasmanian devil	Opossum	Platyapus	Mean expansion of monotocous species	Mean expansion of polytocous species	p-value	
Monotocous specific expansion	<i>EPHA3</i>	<i>EPHA4</i>	7	7	7	6	6	6	5	7	5	7	7	7	5	5	6	5	6.88	5.43	1.10×10 ⁻⁴	
	<i>EPHA5</i>	<i>EPHA6</i>																				
	<i>EPHA7</i>	<i>EPHA8</i>																				
	<i>EPHA10</i>																					
Polytocous specific expansion	<i>REEP1</i>	<i>REEP2</i>	6	5	5	5	6	6	6	5	6	5	6	5	6	6	6	5	5.25	6.00	2.54×10 ⁻³	
	<i>REEP3</i>	<i>REEP4</i>																				
	<i>REEP5</i>	<i>REEP6</i>																				

¹ Monotocous mammals are represented with bold letters.

the oviduct, cervix, and vagina (Gu et al., 2005). Leukocytes interact with several physiological process in the reproductive organs as major effector cells (Brännström and Enskog, 2002). *TMEM214* is connected to *CD22* with protein-protein interaction mediated by *PTPRC* and *LSMI* (Warde-Farley et al., 2010). Two other identified proteins were member *ARMCX3* is an integral membrane protein of the mitochondrial membrane and *LRRC19* is a transmembrane protein receptor (Mou et al., 2009; Chai et al., 2009). Although we did not find direct connection between these two genes and reproductive process, we found an interesting association between membrane proteins and ovulation. Mutations of inner mitochondrial membrane peptidase 2-like (*IMMP2L*) gene induced infertility in female mouse due to defects in folliculogenesis and ovulation, and indirectly induced infertility in male mouse by erectile dysfunction (Lu et al., 2008). While it is possible that *ARMCX3* and *LRRC19* may function in a similar manner to the other membrane proteins associated with infertility, more research is needed to better understand their role in determining litter size.

Gene expansion

Genes under expansion in the monotocous and polytocous species were identified from the 7,508

orthologous gene sets including in-paralogs of leaf species from the phylogenetic topology. The gene family in monotocous species that has a significantly higher average number of genes that have undergone expansion than that in polytocous species and vice versa is shown in Table 1. Unfortunately, the smallest expansion number of the *EPHA* gene family of monotocous species is 6 while the highest expansion number of polytocous species is also 6. Among the 7,508 orthologous gene sets, there was no single gene family for which the lowest expansion number for the target trait (i.e. monotocous) was higher than the highest expansion number of its counterpart trait (i.e. polytocous) or vice versa. Table 2 shows the results of the most significant difference in expansion number between monotocous and polytocous.

In an association study of *EPHA4* polymorphism with swine reproductive traits, it was revealed that the *EPHA4* gene was significantly associated with litter size in pigs. The locus seemed to confer advantages to litter size supporting the results of this study (Fu et al., 2012a). We found that the gene, *EPHA4*, has undergone expansion in monotocous species and this results is supported by previous findings. Taking together the previous results with the gene family identified in this study, the *EPHA4* gene which effects litter size has the possibility of affecting litter

Table 2. Monotocous specific and polytocous specific genes in monotocous and polytocous mammals

Type	Gene family	Total score	Query coverage (%)	Identity (%)	E-value
Monotocous specific orthologous gene sets	<i>APOBEC3A, APOBEC3B, APOBEC3G</i>	387	76	40	9.00×10 ⁻⁸⁶
	<i>ZNF619</i>	826	99	67	0
	<i>TSPAN16</i>	275	94	62	2.00×10 ⁻⁹¹
Polytocous specific orthologous gene sets	<i>Gm6214, Zcchc10</i>	162	41	95	8.00×10 ⁻⁴⁹
	<i>Wfdc6b, Wfdc6a</i>	169	98	57	9.00×10 ⁻⁵³

Monotocous specific and polytocous specific genes are orthologous gene sets which only exist in either monotocous or polytocous species groups. They are mined from the total orthologous gene set information and has <70% query coverage or <70% identity supported by blastp from the NCBI reference gene database.

Genes with <50% query coverage or <50% identity are represented with bold letters.

size in monotocous species. In addition, another gene of the ephrin family, *EPHB2* has also been revealed to have a significant association with litter size in pigs (Fu et al., 2012b). However, our results did not show a significant difference between the two groups. The average expansion was 6.5 for monotocous species and 6.7 for polytocous species (p-value of 0.52). Therefore, among the two genes which affect litter size in the ephrin family, *EPHA4* and *EPHB2*, the results of the analysis suggests that only *EPHA4* might play a role in monotocous species.

In the orthologous gene sets of the 16 species, gene expansion was analyzed among in-paralogs within the leaf of the phylogenetic tree which is the presently existing species. However, if a portion of the current species were analyzed for gene expansion by grouping them under a common ancestor, such as part of the monotocous species (human, chimpanzee, orangutan, and macaque) or polytocous species (cow and dolphin), a higher number of orthologous gene could have been assessed. This is due to the fact that the gene set that arose from duplication in human, chimpanzee, orangutan, and macaque, is out-paralogs to the existing species and so becomes eliminated from the dataset for gene expansion testing.

Monotocous and polytocous specific gene sets

The monotocous species-specific orthologous gene family and polytocous species-specific orthologous gene family were obtained from orthologous information of 319,682 genes of 16 species including in-paralogs. These group-specific genes were searched against the database of the counterpart species to check for genes with high levels of similarity i.e. the monotocous species-specific genes were searched against the polytocous species database and vice versa (Table 2). Candidate genes with over 70% identity and 70% coverage similarity (Martinez et al., 2008) were filtered out leaving 5 gene families. Among these gene families, the *APOBEC* gene family in the monotocous specific gene set, and *Gm6214* and *Zcchc10* families in the polytocous specific gene sets showed <50% identity and <50% coverage, indicating that they do not exist in the counterpart groups.

APOBEC3C is connected to *CLDN4* with protein-protein interaction mediated by *SHBG* (Warde-Farley et al., 2010), which is a trans-membrane protein that might negatively influence fertility rates and is associated with assisted reproduction outcome (Serafini et al., 2009).

ACKNOWLEDGEMENTS

This work was supported by a grant from the Next Generation BioGreen 21 Program (No.PJ009019), Rural Development Administration, Republic of Korea.

AUTHOR CONTRIBUTIONS

Hyeonju Ahn and Kyu-Won Kim carried out analysis and interpretation of data and mainly drafted the manuscript. Hyeon-Jeong Kim and Seoae Cho contributed to the interpretation of the data and the drafting of the manuscript. Heeбал Kim supervised the entire study and helped with the paper.

REFERENCES

- Abe, Y., H. Sinozaki, T. Takagi, T. Minegishi, K. Kokame, K. Kangawa, M. Uesaka, and K. Miyamoto. 2006. Identification of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (TCDD)-inducible genes in human amniotic epithelial cells. *Reprod. Biol. Endocrinol.* 4:1-9.
- Brännström, M. and A. Enskog. 2002. Leukocyte networks and ovulation. *J. Reprod. Immunol.* 57:47-60.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540-552.
- Chai, L., L. Dai, Y. Che, J. Xu, G. Liu, Z. Zhang, and R. Yang. 2009. LRRC19, a novel member of the leucine-rich repeat protein family, activates NF- κ B and induces expression of proinflammatory cytokines. *Biochem. Biophys. Res. Commun.* 388:543-548.
- Drogemüller, C., H. Hamann, and O. Distl. 2001. Candidate gene markers for litter size in different German pig lines. *J. Anim. Sci.* 79:2565-2570.
- Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5:163-166.
- Fu, Y., J. Fu, and A. Wang. 2012a. Association of EphA4 polymorphism with swine reproductive traits and mRNA expression of EphA4 during embryo implantation. *Mol. Biol. Rep.* 39:2689-2696.
- Fu, Y., G. Wang, J. Fu, and A. Wang. 2012b. Association of EphrinB2 gene polymorphism with litter size in pigs. In: *Information Technology and Agricultural Engineering* (Ed. E. Zhu and S. Sambath). *Adv. Intell. Soft Comput.* 134:811-818.
- Galloway, S. M., K. P. McNatty, L. M. Cambridge, M. P. Laitinen, J. L. Juengel, T. S. Jokiranta, R. J. McLaren, K. Luiro, K. G. Dodds, G. W. Montgomery, A. E. Beattie, G. H. Davis, and O. Ritvos. 2000. Mutations in an oocyte-derived growth factor gene (*BMP15*) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat. Genet.* 25:279-283.
- Gu, W., P. Janssens, M. Holland, R. Seamark, and P. Kerr. 2005. Lymphocytes and MHC class II positive cells in the female rabbit reproductive tract before and after ovulation. *Immunol. Cell Biol.* 83:596-606.
- Hedges, S. B., J. Dudley, and S. Kumar. 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22:2971-1972.
- Hubbard, T., D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, and T. Down. 2002. The Ensembl genome database project. *Nucl. Acids Res.* 30:38-41.
- Jennen, D. G., C. Magkoufopoulou, H. B. Ketelslegers, M. H. van Herwijnen, J. C. Kleinjans, and J. H. van Delft. 2010.

- Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol. Sci.* 115:66-79.
- Löytynoja, A. and N. Goldman. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. USA.* 102:10557-10562.
- Lu, B., C. Poirier, T. Gaspar, C. Gratzke, W. Harrison, D. Busija, M. M. Matzuk, K. E. Andersson, P. A. Overbeek, and C. E. Bishop. 2008. A mutation in the inner mitochondrial membrane peptidase 2-like gene (*Immp2l*) affects mitochondrial function and impairs fertility in mice. *Biol. Reprod.* 78:601-610.
- Martinez, D., R. M. Berka, B. Henrissat, M. Saloheimo, M. Arvas, S. E. Baker, J. Chapman, O. Chertkov, P. M. Coutinho, and D. Cullen. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat. Biotechnol.* 26:553-560.
- Mou, Z., A. R. Tapper, and P. D. Gardner. 2009. The armadillo repeat-containing protein, ARM CX3, physically and functionally interacts with the developmental regulatory factor Sox10. *J. Biol. Chem.* 284:13629-13640.
- Pruitt, K. D., T. Tatusova, and D. R. Maglott. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucl. Acids Res.* 35:D61-D65.
- Rothschild, M., C. Jacobson, D. Vaske, C. Tuggle, L. Wang, T. Short, G. Eckardt, S. Sasaki, A. Vincent, D. McLaren, O. Southwood, H. van der Steen, A. Mileham, and G. Plastow. 1996. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Natl. Acad. Sci. USA.* 93:201-205.
- Serafini, P. C., I. Silva, G. D. Smith, E. Motta, A. M. Rocha, and E. C. Baracat. 2009. Endometrial claudin-4 and leukemia inhibitory factor are associated with assisted reproduction outcome. *Reprod. Biol. Endocrinol.* 7:30.
- Stamenkovic, I., D. Sgroi, A. Aruffo, M. S. Sy, and T. Anderson. 1991. The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and alpha 2-6 sialyltransferase, CD75, on B cells. *Cell* 66:1133-1144.
- Toba, K., H. Hanawa, M. Sakaue, K. Yoshida, H. Itoh, J. Tsuchiyama, S. Maruyama, M. Narita, M. Takahashi, K. Watanabe, and Y. Aizawa. 2003. Fc epsilon RI and CD22 mRNA are expressed in early B-lineage and myeloid leukemia cell lines. *Leuk Res.* 27:173-182.
- Vincent, A. L. 1997. The prolactin receptor gene is associated with increased litter size in pigs. *Swine Res. Rep.* http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1010&context=swinereports_1997. Accessed January 23, 2014.
- Warde-Farley, D., S. L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, and C. T. Lopes. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucl. Acids Res.* 38:W214-W220.
- Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586-1591.