De novo oviduct transcriptome of the moor frog *Rana arvalis*: a quest for maternal effect candidate genes

Longfei Shu*[1], Jie Qiu[2] and Katja ~~Räsänen~~Räsänen[1]

Affiliations:

~~Eawag~~[1]Eawag, Department of Aquatic Ecology, Switzerland and ETH Zürich, Institute of Integrative Biology, Switzerland

[2]Institutue of Crop Science & Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

Corresponding author:

Dr. Longfei Shu

Eawag, Department of Aquatic Ecology, Switzerland and ETH Zürich, Institute of Integrative Biology, CH-8600 Dübendorf, Switzerland;

Email: longfei.shu@wustl.edu; Phone: +1-314-445-8037.

*__Current address:__ Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA

~~**Keywords**: Amphibian; Egg coat; Glycosylation; Lake acidification; Maternal effect genes; Moor frog; RNA Seq; Rana arvalis; Transcriptome~~

~~**Running title:** First transcriptome of the moor frog~~

Abstract

Maternal effects can substantially affect ecological and evolutionary processes in natural populations. However, as they often are environmentally induced, establishing their genetic basis is challenging. One important, but largely neglected, source of maternal effects are egg coats (i.e. the maternally derived extracellular matrix that surrounds the embryo). In *Rana arvalis*, ~~an emerging ecological and evolutionary amphibian model,~~ the gelatinous egg coats (i.e. egg jelly) are produced in the mother's oviduct and consist primarily of highly glycosylated mucin type O-glycans. These O-glycans affect jelly water balance and, subsequently, strongly contribute to adaptive divergence in embryonic acid tolerance.

To identify candidate genes for maternal effects, we conducted RNAseq transcriptomics on oviduct samples from seven *R. arvalis* females, representing the full range of within and among population variation in embryonic acid stress tolerance across our study populations. *De novo* sequencing of these oviduct transcriptomes detected 124 071 unigenes and functional annotation analyses identified a total of 57 839 unigenes, of which several identified genes likely code for variation in egg jelly coats. These belonged to two main groups: mucin type core protein genes and five different types of glycosylation genes. We further predict 26 711 gene-linked microsatellite (SSRs) and 231 274 single nucleotide (SNPs) polymorphisms.

Our study provides the first set of genomic resources for *R. arvalis*, an emerging model system for the studies of ecology and evolution in natural populations, and will aid in gaining insight to the genetic architecture of egg coat mediated maternal effects.

Background

Understanding evolutionary processes of natural populations necessitates a good understanding of the genetic architecture of trait variation in an ecologically relevant context (e.g. ~~(Houle et al. 2010; Mitchell-Olds et al. 2007; Nadeau & Jiggins 2010)~~(Houle et al. 2010; Mitchell-Olds et al. 2007; Nadeau & Jiggins 2010)). Maternal effects (the effects of ~~mother~~a mother's environment and phenotype on offspring performance) are an important source of phenotypic variation and often under strong natural selection (reviewed in

~~(Mousseau & Fox 1998; Räsänen & Kruuk 2007)~~(Mousseau & Fox 1998; Räsänen & Kruuk 2007)). Maternal effects can influence speed and direction of evolution as well as facilitate local adaptation ~~(Hangartner et al. 2012; Räsänen & Kruuk 2007; Shu et al. 2016; Wolf et al. 1998)~~(Hangartner et al. 2012; Räsänen & Kruuk 2007; Shu et al. 2016; Wolf et al. 1998). However, as maternal effects are ~~often~~typically at least partially environmentally induced, their genetic architecture is ~~still~~ poorly understood.

Maternal effects can arise through various mechanisms, most commonly acknowledged through variation in egg size and content (reviewed in (Bernardo 1996; ~~Mousseau & Fox 1998~~Mousseau & Fox 1998). A much less well acknowledged but important source of maternal effects are so called egg coats ~~(Shu et al. 2015b)~~(Shu et al. 2015b)~~. Egg coats~~. Egg coats (present in all sexually reproducing animals, as well as many asexual metazoans) are maternally derived, extracellular structures that consist of multiple functionally and structurally different layers (reviewed in ~~(Menkhorst & Selwood 2008; Shu et al. 2015b; Wong & Wessel 2006)~~(Menkhorst & Selwood 2008; Shu et al. 2015b; Wong & Wessel 2006)). These structures have many key functions, ranging from fertilization to embryonic protection ~~(Shu et al. 2015b)~~(Shu et al. 2015b). The genetic basis of the innermost, so called oocyte coats, has been studied in several model systems, such as sea urchins and the abalone (Claw & Swanson 2012; ~~Palumbi 2009~~Palumbi 2009), the moth *Bombyx mori* ~~(Papantonis et al. 2015)~~(Papantonis et al. 2015) and the frog *Xenopus laevis* ~~(Hedrick 2008; Shu et al. 2015b)~~(Hedrick 2008; Shu et al. 2015b). However, the genetic basis of ~~the outer~~outermost gelatinous egg coats~~ – that surround embryos of many~~, often found in taxa where embryos develop externally in the surrounding environment (e.g. various invertebrates, fish and amphibians~~)~~), is to date effectively unstudied (reviewed in (Shu et al. 2015b)).

~~Recently, we showed that in~~In the moor frog *Rana arvalis*, jelly coat mediated maternal effects have driven adaptive divergence in embryonic acid stress tolerance ~~(Hangartner et al. 2011; Persson et al. 2007; Räsänen et al. 2003; Shu et al. 2015a; Shu et al. 2016; Shu et al. 2015c)~~(Hangartner et al. 2011; Persson et al. 2007; Räsänen et al. 2003b; Shu et al. 2015a; Shu et al. 2016; Shu et al. 2015c). ~~Jelly~~This divergence is primarily due to glycoproteins of the jelly coats, which influence water balance and, subsequently, embryonic survival in acidic conditions ~~(Shu et al. 2015a)~~(Shu et al. 2015a). Egg jelly ~~coats consist of~~ glycoproteins~~, with~~ are complex glycan structures attached to a protein backbone ~~(Hedrick & Nishihara 1991)~~(Hedrick & Nishihara 1991) and are highly species specific (Coppin et al. 1999; Delplace et al. 2002~~; Strecker et al. 2003)~~; Strecker et al. 2003). Given this complexity, the

Field Code Changed

jelly coats can be coded by multiple genes that regulate – for instance – the protein backbone and the activities of enzymes that impact how different branches are attached ~~(Shu et al. 2015b)~~(Shu et al. 2015b). This makes identifying the genetic basis of jelly coat variability highly challenging – particularly so in natural populations not amenable to experimental cross-generational rearing or genetic manipulations. Here next generation sequencing tools, in particular RNA-seq, can be helpful ~~(Todd et al. 2016; Wang et al. 2009)~~(Todd et al. 2016; Wang et al. 2009). Here we ~~apply~~applied *de novo* transcriptomics to get first insight to jelly coat genes in ~~the moor frog (~~*~~Rana~~R. arvalis*~~)~~. We ~~conduct~~conducted tissue specific RNA-seq ~~on the oviducts~~ of *R. arvalis* oviducts, where jelly coat biosynthesis takes place ~~(Hedrick & Nishihara 1991)~~(Hedrick & Nishihara 1991).

Despite the increasing numbers of genomes being sequenced, the genomic resources for amphibians ~~remain depauperate as~~ are sparse and only ~~two~~a handful of amphibian genomes ~~(*Xenopus tropicalis* and *Nanorana perkeri*)~~ have been sequenced so far ~~(Hellsten et al. 2010; Sun et al. 2015)~~(Elewa et al. 2017; Hammond et al. 2017; Hellsten et al. 2010; Nowoshilow et al. 2018; Sun et al. 2015). This is an important shortcoming, given that amphibians are common model systems for a range of ecotoxicological ~~(Helbing 2012)~~(Helbing 2012), ecological and evolutionary studies, the latter ranging from spatial patterns of phenotypic and genetic divergence (Egea-Serrano et al. 2014~~; Richter-Boix et al. 2011; Van Buskirk & Arioli 2005)~~; Richter-Boix et al. 2011; Van Buskirk & Arioli 2005) to developmental plasticity ~~(Gomez-Mestre et al. 2006)~~(Gomez-Mestre et al. 2006) and adaptive maternal effects ~~(Mousseau & Fox 1998; Räsänen et al. 2003; Shu et al. 2016; Shu et al. 2015c)~~(Mousseau & Fox 1998; Räsänen et al. 2003b; Shu et al. 2016; Shu et al. 2015c).

To bridge this knowledge gap – and as a first step to identify candidate genes for egg coat mediated maternal effects - we here apply tissue specific transcriptomics on *R. arvalis* oviducts. In doing so, we complement recent transcriptomics studies on other Ranid species (Birol et al. 2015~~; Helbing 2012; Price et al. 2015; Qiao et al. 2013; Zhang et al. 2013)~~; Helbing 2012; Price et al. 2015; Qiao et al. 2013; Robertson & Cornman 2014; Zhang et al. 2013) and increase availability of genomics resources for amphibians in general. As our specific interest is in identifying putative genes underlying egg jelly coat variation, we collected samples from the oviduct on seven *R. arvalis* individuals ~~bracketing~~encompassing the full range of embryonic acid tolerance variation among and within our study populations ~~(Hangartner et al. 2011; Shu et al. 2015a; Shu et al. 2015c)~~(Hangartner et al. 2011; Shu et al. 2015a; Shu et al. 2015c)~~. We implemented *de novo* transcriptome sequencing and~~. We focused particularly on identifying genes related to biosynthesis of mucin type O-linked glycans. In addition, ~~we provide the first set of transcriptomic and genome wide molecular markers for *R. arvalis*.~~this study provides the first transcriptomes for *R. arvalis*, and add more genome wide markers in addition to existing resources (Brelsford et al. 2017).

Materials and Methods

Study system

R. arvalis is an anuran amphibian, broadly distributed across the western Palearctic (Glandt 2006). Individuals from three *R. arvalis* populations (Tottajärn, [T~~,~~], Bergsjö, [B] and Stubberud, [S~~)~~]) breeding in permanent ponds along an acidification gradient in southwestern Sweden were used in this study ~~(Hangartner et al. 2011; Shu et al. 2015a; Shu et al. 2016; Shu et al. 2015b; Shu et al. 2015c)~~(Hangartner et al. 2011; Shu et al. 2015a; Shu et al. 2016; Shu et al. 2015b; Shu et al. 2015c)~~.~~. These ponds differ in pH due to a mix of natural acidification processes, variation in bedrock buffering capacity and acid rain (Räsänen et al. 2003a; Räsänen et al. 2003b). These populations were chosen ~~as~~because, based on former common garden experiments ((Hangartner et al. 2012; Shu et al. 2015c), they represent the extreme ends of adaptive divergence ~~along an acidification gradient in~~ in embryonic acid stress tolerance (i.e. [S] most acid sensitive; [B] intermediate tolerance; [T] most acid tolerant~~; (Hangartner et al. 2012; Shu et al. 2015c)).~~). The pH in these ponds ranges from highly acidic (pH 4, site [T~~)~~]) to intermediate (pH 6, site [B~~)~~]) to neutral (pH 7.5, [S~~).~~]).

During the breeding season of 2012, females in breeding condition were collected and transported to the laboratory at Uppsala University (59°50`N, 17°50`E). The females were maintained in containers with moist *Sphagnum* moss (antibacterial medium) in a climate chamber at 2- 4°C until artificial crosses (see below) and RNA sampling were conducted a few days later. The females used for this transcriptomics study consist of a subset (out of 7 to 10 females/population) ~~used in artificial crosses~~of females that were artificially crossed to establish acid tolerance of embryos in a common garden laboratory experiment and to study variation in macromolecular composition of egg jelly coats (for details see ~~(Shu et al. 2016; Shu et al. 2015c)~~(Shu et al. 2016; Shu et al. 2015c)). The experiments were conducted under permissions from the Västra Götaland county board (collection permit: Dnr 522-6666-2011) and the Ethical committee (Dnr C65/11) for animal experiments in Uppsala County.

Sampling and RNA extraction

To bracket a broad range of genetic backgrounds, ~~the individuals chosen (~~a total ~~N = 6) for RNA analyses~~of six females were ~~selected~~chosen so as to represent females that produced the most acid tolerant (highest embryonic survival at acidic pH in a common garden experiment) and most acid sensitive ~~clutch~~(lowest embryonic survival at acidic pH in a common garden experiment) clutches both among and within each of the three populations ~~in Shu *et al.* (Shu et al. 2015c).~~(Shu et al. 2016). In addition, we included a sample from one female that had not yet fully ovulated, hence providing a reference for gene expression at an earlier stage of egg coat production~~.~~ (Supplementary material). For each female, RNA of the whole oviduct (i.e. specific tissue where egg jelly is produced; ~~(Hedrick & Nishihara 1991)~~(Hedrick & Nishihara 1991)) was collected~~.~~ and stored in RNA later at -20 °C until extraction. RNA extraction was conducted using TRIzol (Life Technologies) according to the manufacturer's protocol, followed by DNase (Qiagen) treatment to eliminate potential genomic DNA contamination. Both the concentration and integrity of the RNA samples for transcriptomic analyses were evaluated with the Angilent 2100 Bioanalyzer. All samples had an RNA integrity value (RIN) > 8 and were, hence, used to construct the cDNA libraries ~~(Schroeder et al. 2006)~~(Schroeder et al. 2006).

cDNA library construction and sequencing

cDNA libraries were generated using the TruSeq RNA-Seq Sample Prep kit according to the manufacturer's protocol (Illumina Inc., San Diego, CA, USA). Briefly, magnetic beads with Oligo (dT) were used to isolate the poly(A)+ mRNA. Fragmentation buffer was added in the presence of divalent cations to break the mRNA into short fragments of approximately 200 bp. Short fragments were purified with QiaQuick PCR extraction kit, followed by end reparation, adding poly(A) and sequencing adapters. The suitable fragments were selected for the PCR amplification as templates. In total, seven cDNA libraries were constructed and sequenced using the Illumina HiSeq^{TM} 2000 (90 bp paired-end reads). The sequencing reactions were conducted at the Beijing Genomics Institute (BGI), Shenzhen.

Assembly and annotation

Raw reads were filtered to remove reads containing adaptors, reads with unknown nucleotides greater than 5% and low-quality reads with more than 20% bases of quality value ≤ 10 (filter_fq, BGI). Only clean reads were used in the following analyses. Transcriptome *de novo* assembly was carried out using the assembly program Trinity ~~(Grabherr et al. 2011)~~(Grabherr et al. 2011). Briefly, the software first combined reads of certain lengths of overlap to form contigs. Subsequently, the reads were mapped back to the contigs, which were connected until extension proceeded on neither end ~~(Grabherr et al. 2011)~~(Grabherr et al. 2011). After removing any redundancy, the resulted sequences were defined as unigenes. Software and their parameters ~~are listed below~~were: Trinity (--seqType fq --min_contig_length 100; --min_glue 4 --group_pairs_distance 250; --path_reinforcement_distance 85 --min_kmer_cov 4); TGICL (-l 40 -c 10 -v 20); Phrap (-repeat_stringency 0.95 -minmatch 35 -minscore 35).

Unigene annotation provides information of expression and functional annotation of Unigene. Unigene sequences were first annotated using blastx against the database Nr (http://www.ncbi.nlm.nih.gov/), with a cut-off E-value of 1e-5. To acquire a more comprehensive annotation, Unigene sequences were also aligned to the protein databases Swiss-Prot, ~~KEGG~~Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) (1e-5) by blastx. In order to predict and classify the possible functions of the unigenes, they were additionally annotated to Cluster of Orthologous Group (COG), which classifies orthologous gene products (Tatusov *et al.* 2003). Software and their parameters ~~are listed below~~were: BLAST (-F F -e 1e-5 -p blastn; -F F -e 1e-5 -p blastx); Blast2GO (Default); Path_finder (Default).

Protein Coding Sequence Prediction (CDS)

Unigenes were aligned by blastx (*e* value < 0.00001) to protein databases in the priority order of NR, Swiss-Prot, KEGG and COG. Unigenes aligned to a higher priority database were not aligned to lower priority database. Proteins with highest ranks in the blast results were used to decide the coding region sequences of the Unigenes. The coding region sequences were subsequently translated into amino acid sequences with the standard codon table. This produced both the nucleotide sequences (5'->3') and amino acid sequences of the Unigene coding region. Unigenes that could not be aligned to any database were scanned by ESTScan ~~(Iseli et al. 1999)~~(Iseli et al. 1999), producing nucleotide sequence (5'->3') direction and amino sequence of the predicted coding region.

Metabolic pathway analysis

To investigate the functions of the unigenes in metabolic process, we acquired pathway annotation by mapping the unigenes to ~~Gene Ontology (~~GO~~)~~ and ~~Kyoto Encyclopedia of Genes and Genomes (~~KEGG~~)~~ database ~~(Kanehisa et al. 2008)~~(Kanehisa et al. 2008). ~~Gene Ontology (~~GO~~)~~ database is a relational database, which has three ontologies: molecular function, cellular component and biological process. KEGG is a database for understanding high-level functions and utilities of the biological system. It is able to analyze gene products during metabolism and cellular processes and allowed us to get pathway annotation for unigenes. We used the Blast2GO program with default setting to acquire GO and KEGG annotation of Unigenes ~~(Tarazona et al. 2011)~~(Tarazona et al. 2011).

Identification of candidate genes for maternal effects

Based on prior work, jelly coats of *R. arvalis* and other amphibians ~~As~~(Coppin et al. 1999; Guerardel et al. 2000; Strecker et al. 2003) mostly consist of mucin type O-glycans, making genes related to mucins and to O−glycan biosynthesis particularly relevant as maternal effect genes. Therefore we considered as candidate genes related to *R. arvalis* egg jelly coats ~~(Shu et al. 2015b)~~(Shu et al. 2015b)~~, we considered~~ unigenes that i) were expressed in all individuals, ii) were mapped to the category of “extracellular matrix” in GO and KEGG annotations and iii) were involved in the glycosylation process~~ of the egg jelly. ~~. ~~Based on prior work, jelly coats of *R. arvalis* and other amphibians (Coppin et al. 1999; Guerardel et al. 2000; Strecker et al. 2003) mostly consist of mucin type O-glycans, making genes related to mucins and to O−glycan biosynthesis particularly relevant as maternal effect genes.~~

~~-~~***Preliminary differential expression analysis***

Unigene expression was calculated using the FPKM (Fragments Per kb per Million reads), which can eliminate the influence of different gene length and sequencing level on the calculation of gene expression ~~(Mortazavi et al. 2008)~~(Mortazavi et al. 2008). Group comparison between populations was performed using the R Bioconductor package NOISeq, which is a data-adaptive and nonparametric method ~~(Tarazona et al. 2011)~~(Tarazona et al. 2011)~~.~~. We defined Probability ≥ 0.8 and the absolute value of Log2Ratio ≥ 1 as the threshold to judge the significance of gene expression difference between populations. KEGG enrichment analysis was performed with a Fisher’s exact test in Blast2GO ~~(Tarazona et al. 2011)~~(Tarazona et al. 2011). Pathway enrichment analysis identifies significantly enriched metabolic pathways or signal transduction pathways in differential expressed genes (DEGs). After correction for multiple testing, we chose pathways with Q value ≤ 0.05 as significantly enriched in DEGs ~~(Gotz et al. 2008; Tarazona et al. 2011)~~(Gotz et al. 2008; Tarazona et al. 2011).

Formatted: Space After: 10 pt

Genetic marker resources

To aid population genomic analyses for taxa were no genome is available, RNA-seq data can be used to identify genomic markers, such as *Simple sequence repeats (*SSR) and *Single nucleotide polymorphisms (*SNPs), in transcribed regions ~~(Lopez-Maestre et al. 2016)~~(Lopez-Maestre et al. 2016). ~~Here we~~We therefore used the oviduct RNA-seq data to predict SSR and SNP markers, ~~to be~~which allows later ~~validated for~~validation of targeted questions ~~(Lopez-Maestre et al. 2016)~~(Lopez-Maestre et al. 2016)~~.~~, such as allelic frequency variation of the candidate genes in the source populations.

SSRs were identified in the final assembly with the software MicroSAtellite (MISA, http://pgrc.ipk-gatersleben.de/misa/) using unigenes as the reference. Assembled contigs were scanned for SNPs with SNP detection software SOAPsnp ~~(Li et al. 2008)~~(Li et al. 2008). The program can assemble consensus sequence for the genome of a newly sequenced individual based on the alignment of the raw sequencing reads on the unigenes. The program calculated the likelihood of each genotype at each site based on the alignment of short reads to a unigene, set together with the corresponding sequencing quality scores. It then infers the genotype with highest posterior probability at each site based on Bayes' theorem (the reverse probability model ~~(Li et al. 2008)~~(Li et al. 2008)). Therefore, the intrinsic bias or errors that are common in Illumina sequencing data have been taken into account and the quality values for use in inferring consensus sequence have been recalibrated. Software and their parameters ~~are listed below~~were: SOAPsnp (Release 1.03, http://soap.genomics.org.cn/soapsnp.html, -u t -Q i -L 90).

Additional analyses

To test whether embryonic acid tolerance is related to the transcriptional activity, we performed correlation analysis between the acid tolerance of each individual and the expression of each unigene (FPKM) by PYTHON. A strict correlation threshold was applied (Pearson correlation, $|r| > 0.9$ and p-value < 0.01). To test how much of the fluke genomes our *R. arvalis* dataset cover, we downloaded the genomes of 3 flukes (*Clonorchis sinensis, Schistosoma mansoni, Schistosoma japonicum*) from the wormbase (https://parasite.wormbase.org/), and BLAST them (BLATP, E-value < 1e-7) against our unigenes dataset. To examine whether there are any hits from our database to the recently developed transcriptomic resources for the green frog *Rana clamitans* and chorus frog *Pseudacris regilla* (Robertson & Cornman 2014), we downloaded the assembled transcriptomes from NCBI BioProjects (PRJNA162931 and PRJNA163143), and used them as query to blast against our Unigene database (BLASTX, E-value < 1e-7).

Results

Sequencing and assembly

In total, 53 330 025 420 bp bases were generated from the *R. arvalis* oviduct transcriptome. Total clean reads of the seven cDNA libraries ranged from 81 166 804 to 87 485 924 (Table 1), with an average GC content of 44.95%. In the final assembly, ~~69 987~~87 401 to 112 136 unigenes were detected across the ~~seven~~six cDNA libraries (Table 2). Interestingly, for the ~~immature~~one female ~~has less~~that had not yet ovulated fewer unigenes (69 987) were detected than ~~mature~~for the remaining females~~.~~ (that had ovulated). When the cDNA libraries were pooled, a total of 124 071 unigenes were detected, with an N50 of 1, 212 bp and a total length of 90.3 Mb. The average length of the unigenes was 728 bp (~~the size distribution of the unigenes is shown in Fig.~~Fig. 1).

Transcriptome annotation

The E-value distributions of the unigenes showed that approximately 60% of the unigenes had very strong homogeny (< 1e-30) with the Nr database, while the rest ranged from 1e-5 to 1e-30 (Fig. 2A). Thirty percent of the unigene sequences had over 80% similarity with the Nr database, while the similarity of the remaining 70% of sequences ranged from 17% to 80% (Fig. 2B). The *R. arvalis* sequences matched best with *Xenopus (Silurana) tropicalis* (43.4%), followed by *Xenopus laevis* (13.1 %, Fig. 2C) and the liver fluke *Clonorchis sinensis* (12.1 % Fig. 2C).

Functional classes were successfully annotated for 13 501 unigenes using COG (Fig. 3, 4). BLASTX against Swiss-Prot, KEGG, NT and GO database resulted in 37 262, 31 405, 39 138 and 24 452 hits, respectively (Fig. 4). Altogether, 57 839 *R. arvalis* unigenes (46.6% of all 124 071 unigenes) had significant matches with existing databases (Fig. 4). Among the 124 071 unigenes, 48 850 (39.4%) were predicted as Protein coding sequences (CDS). Of these, 44 809 unigenes were aligned using existing databases, while 4 041 unigenes ~~which~~that could not be annotated with any database were predicted by EST Scan. The length distribution of the CDS protein sequences is available in the Supporting information.

Functional pathway annotation

A total of 24 452 unigenes (42.27% of all annotated unigenes) were successfully categorized into 60 GO functional groups (Fig. 5). These were classified to three major categories: biological processes (23), cellular components (18) and molecular functions (19) (Fig. 5). A total of 31 405 unigenes (54.29% of all annotated unigenes) were annotated in KEGG, and assigned to 259 known KEGG pathways (Table S2). The highly enriched pathways included: metabolic pathways (3796, 12.09%), purine metabolism (1910, 6.08%), regulation of actin cytoskeleton (1202, 3.83%), focal adhesion (1161, 3.7%) and calcium signaling pathway (1159, 3.69%).

Candidate genes for egg jelly coat

We identified two groups of candidate genes likely related to egg jelly coats: 1) core protein genes (Table 3) and 2) protein glycosylation genes (Table 4). The major components of egg jelly core protein (extracellular matrix fiber, ECM) are ~~mucin and collagen (Hay~~mucins (Hedrick 1991), of which 13 and 11 different types, respectively, were detected in our dataset (Table 3). Within the Mucin gene family, Mucin-5AC, Mucin-5B and Mucin-2 were very highly expressed (Table 3) and the most abundant transcripts of all unigenes. In addition, previous evidence suggests that the Mucin-5 gene is expressed exclusively in the *X. tropicalis* oviduct, making Mucin-5 genes likely candidates for egg jelly coat genes ~~(Lang et al. 2016)~~(Lang et al. 2016). Several other minor components were also identified (Dermatopontin, Fibulin-5 and Fibrinogen-like protein 1, Decorin, EMILIN-1, Fibrillin-1, Fibronectin and Laminin) (Table 3).

In addition to the candidate core protein genes, five major biosynthesis pathways involved in protein glycosylation were identified based on KEGG: Mucin type O−Glycans, Other types of O−glycans, Heparan sulfate, Chondroitin sulfate and Keratan sulfate (Table 4). Of these, Mucin type O−Glycan genes are the most likely candidates for egg jelly coat glycosylation (Coppin et al. 1999~~; Lang et al. 2016)~~; Lang et al. 2016) (Fig. 6). Of the eight types of Mucin type O−Glycans, only genes related to cores 1, 2, 3, 4 and 6 Mucin type O−Glycan biosynthesis were detected, while cores 5, 7 and 8 were not expressed in *R. arvalis* (Fig. 6).

Preliminary differential expression analysis

Differentially expressed genes (i.e. DEGs) were identified by pairwise comparisons between each population pair ~~(~~([S] vs [T~~,~~], [S] vs [B~~,~~], [T] vs [B~~,~~]) of six of the ~~immature~~females (the not ovulated female was excluded~~).~~ from these analyses). Overall, 4457, 4198 and 3691 differentially expressed genes were identified between [S] vs [T~~,~~], [S] vs [B] and [T] vs [B~~,~~], respectively~~.~~ (Figure S3). In general, [T] and [B] individuals had much lower gene expression levels compared to S individuals~~.~~ (Figure S3). For instance, 3397 and 3388 genes were down-regulated in the [T] and [B] individuals compared to the [S] individuals, while only 1060 and

810 genes were up-regulated in [T] individuals compared to the [S] individuals~~.~~ (Figure S3). The smallest number of differentially~~-~~ expressed genes occurred between the [T] and [B] individuals, where 2256 were up-regulated and 1435 were down-regulated in [T] versus [B] individuals~~.~~ (Figure S3).

Enrichment of KEGG pathway in the differentially expressed unigenes was assessed using a Fisher's exact test (FDR < 0.05). The FDR analyses identified 37 and 49 significantly enriched KEGG pathways in [T] (Table S4) and [B] (Table S3) compared to [S] individuals, whereas only two were identified between [T] and [B] (Table S5). Again, the suggested expression profiles of [T] and [B] individuals were more divergent compared to that of [S] individuals, while differentiation between the [T] *vs*~~.~~ .[B] was very small. The Ribosome (KO03010) and Oxidative phosphorylation (KO00190) were the most enriched pathways in both [S] *vs*. [T] and [S] *vs*. [B] comparisons (Table S3, S4). This indicated that [T] and [B] females had, in general, lower rates of energy production as well as protein biosynthesis.

Because we are able to rank each individual frog's degree of acid tolerance based on our previous work (Shu et al. 2016), we can test whether a frog's acid tolerance is related to the transcriptional activity. The correlation analyses found 292 unigenes that were significantly correlated with acid tolerance. These genes have diverse functions ranging from signaling to protein biosynthesis, which could be suitable targets for future functional work (SI2). However, it is important to note that to really validate their relevance, future studies are needed. The information of candidate genes were listed in Supplementary Information 2 (SI2).

SSRs and SNPs

A total of 26 711 SSRs were identified across all unigenes, in which Mono-nucleotide repeats (19 215), Di-nucleotide repeats (5 050) and Tri-nucleotide repeats (2 002) were the most abundant SSR motif classes, and Quad-nucleotide repeat (345), Penta-nucleotide repeats (79) and Hexa-nucleotide repeats (20) were detected at much lower frequencies (Fig. S1, Table S1).

A total of 231 274 SNPs were predicted across all unigenes. The number of SNPs in seven individual cDNA library ranged from 63 354 to 86 608 (Fig. S2). The average Ts/Tv ratio (the numbers of transitions and transversions at the SNP sites) of all SNPs was ca. 1.75.

The SSR or SNP data were not analyzed in detail for segregation among populations – as sample size is small. However, as a first step we compared the homologous SNPs differentiated between population [S] (most acid sensitive) and population [T] (most acid tolerant) that differ in acid tolerance. We found 420 candidate SNPs that could be suitable targets for future functional work. The results are listed in Supplementary Information 3 (SI3).

Discussion

Understanding the genetic architecture of maternal effects is important for our ability to understand the relative contribution of genetic *versus* plastic effects in organismal phenotypes and the eco-evolutionary processes of natural populations ~~(Mousseau & Fox 1998; Räsänen & Kruuk 2007)~~(Mousseau & Fox 1998; Räsänen & Kruuk 2007). Yet given that maternal effects are typically strongly environmentally induced, and that the phenotypic variance of offspring is produced in the mother but their fitness consequences ~~are~~ expressed in the offspring, establishing their genetic basis is challenging – particularly in non-model systems. We therefore applied RNA-seq analyses as a first step towards identifying candidate genes for maternal effects that are mediated through the gelatinous egg coats (i.e. egg jelly) in *R. arvalis* ~~(Shu et al. 2015b)~~(Shu et al. 2015b). We next provide first an overview of the general genomic variation revealed, followed by a discussion of putative candidate genes for maternal effects.

General genomic aspects

We characterized 124 071 unigenes from the *R. arvalis* transcriptome, and successfully annotated 57 839 of them (46.6%). These unigenes most closely matched *Xenopus (Silurana) tropicalis* (43.4%), followed by *X. laevis* (13.1%) and the liver fluke *Clonorchis sinensis* (12.1%) (Fig. 2C). ~~This fluke does not exist in areas where our *R arvalis* samples originate from and, hence, this relatively high match to the lung fluke is most likely due to poor annotation of amphibian genomes (as only two amphibian genomes sequenced thus far). It is also possible that genes expressed in the amphibian oviduct may in part orthologous with that of liver fluke, or *R. arvalis* may have been infested by other genetically related parasites (e.g. lung flukes are common (Bolek & Janovy 2007)).~~It is at the first look somewhat surprising that a substantial portion of the unigenes matched to a trematode parasites (Platyhelminthes) rather than any vertebrate genomes. Given that the *Xenopus* genomes are well annotated, it is possible that we have produced transcriptomes of two taxa (an approach that was recently used to detect endoparasites (Santos et al. 2018). We therefore performed additional analyses to test how much of the fluke genomes our *R. arvalis* dataset covers. We found that the percentage of genes in fluke genomes that have hits in our dataset was very high (*C. sinensis*: 61.4%; *S. mansoni*: 72.3%; *S. japonicum*: 58.0%), indicating that we indeed likely sequences parasites in the oviduct wild moor frogs. Trematode parasites are common in amphibians (Sears et al. 2012), but to our knowledge nothing is known of the oviduct parasites in *R. arvalis*. Further studies are clearly needed to confirm this finding.

~~More than half of the unigenes in our database could not be annotated with any existing databases. This mostly reflects the lack of genomics resources in *R. arvalis* or any closely related frog species, but also in amphibians as a whole.~~ Although numerous genomes are now available across a wide range of taxa, only ~~two~~a handful of amphibian genomes ~~(*X. tropicalis* and *Nanorana perkeri*)~~ have been sequenced thus far ~~(Hellsten et al. 2010; Sun et al. 2015)~~(Elewa et al. 2017; Hammond et al. 2017; Hellsten et al. 2010; Nowoshilow et al. 2018; Sun et al. 2015). This lack of amphibian genomes is mostly due to their large genome sizes (due to large proportion of repeat sequences ~~(Sun et al. 2015)~~(Sun et al. 2015)). ~~Clearly~~Yet clearly more work on amphibian genomics is needed ~~given that~~as 1) amphibians are the most ancient class of land-dwelling vertebrates and their genomic resources are essential for understanding vertebrate development and evolution; 2) the understanding of evolutionary processes of amphibian populations would greatly be facilitated by studies on the genomic architecture of trait variation, and 3) because conservation of amphibians that are under serious global decline ~~(Hof et al. 2011; Stuart et al. 2004)~~(Hof et al. 2011; Stuart et al. 2004) could thus be facilitated (Calboli et al. 2011). The ~~transcriptome~~recently sequenced transcriptomes of *R. arvalis*, (this study) and other ~~ranids~~Ranidae (Birol et al. 2015~~; Helbing 2012; Price et al. 2015; Qiao et al. 2013; Zhang et al. 2013)~~; Helbing 2012; Price et al. 2015; Qiao et al. 2013; Robertson & Cornman 2014; Zhang et al. 2013) ~~recently sequenced will aid in developing amphibian genomics resources.~~ will aid in developing amphibian genomics resources. As we used only oviduct tissue, while most other studies use other tissues (e.g. liver, skin), a comparison of oviduct-specific transcriptomes with transcriptomes from other tissues are needed to reliably identify oviduct-specific genes or expression profiles. For instance, when compared with the recently developed transcriptomic resources for the green frog *Rana clamitans* and chorus frog *Pseudacris regilla* (Robertson & Cornman 2014), the percentage of genes in our dataset that have hits in their dataset are 40.61% (20 402/50 238) and 40.62% (19 583/48 213) respectively, indicating a considerable proportion of candidate oviduct-specific genes. Furthermore, the large number of potentially amplifiable SSRs and SNP markers detected in this study (Supporting information) represent an important resource

Formatted: Font: Italic

for applications in population genetics and for the detection of interesting functional genetic variants ~~(Li et al. 2008; Morin et al. 2004; Schunter et al. 2014)~~(Li et al. 2008; Morin et al. 2004; Schunter et al. 2014).

Candidate genes for maternal effects: inside the egg jelly coat

Although there is substantial potential for a genetic basis ~~of~~ in maternal effects ~~(Räsänen & Kruuk 2007)~~(Räsänen & Kruuk 2007), most studies aiming to identify maternal effect genes have focused on their role in early embryonic development *per se* ~~(Tong et al. 2000; Wu et al. 2003)~~(Tong et al. 2000; Wu et al. 2003). To what extent maternal effect genes contribute to adaptive divergence of local populations and response to natural selection at early life stages is therefore still largely unknown.

We identified two groups of candidate maternal effect genes based on their role in egg jelly biosynthesis: mucin core protein genes and protein glycosylation genes. The mucin genes and ~~several~~ O-linked glycosylation genes are particularly likely candidates, given that amphibian egg jelly - including that of *R. arvalis* (Coppin et al. 1999a) - mostly consists of mucin type O-linked glycoproteins (Coppin et al. 1999~~; Guerardel et al. 2000; Strecker et al. 2003);~~; Guerardel et al. 2000; Strecker et al. 2003). Evidence for high mucin gene expression in the oviduct has been found previously in *Rana chensinensis* ~~(Zhang et al. 2013)~~(Zhang et al. 2013) and ~~*Xenopus*~~*X. tropicalis* ~~(Lang et al. 2016)~~(Lang et al. 2016). Whereas Mucin 2 is expressed in various tissues, Mucin 5AC was reported to be exclusively expressed in the oviduct of *X. tropicalis*. Although Mucins, particularly Mucin 5, are likely candidates for the core protein in jelly coats, the role of different genes contributing to the glycosylation of the jelly coat proteins are much more difficult to infer. For instance, the egg jelly coat of *R. arvalis* consists of at least 19 different glycan building blocks (Coppin et al. 1999) and ~~preliminary~~our own analyses indicate high among individual polymorphism in the macromolecular composition of *R. arvalis* jelly ~~is high~~overall (Shu *et al.* 2016~~,~~) as well as in glycans (Shu, Suter and Räsänen, unpublished data). Given the complexity of O-linked glycosylation it is ~~therefore~~ not surprising that multiple glycosylation genes are highly expressed in the *R. arvalis* oviduct - and the genetic basis of jelly coat mediated adaptive maternal effects could be complex.

In general, the genes representing likely candidate genes for egg jelly coat formation ~~showed~~did show a different expression pattern compared to the global profile (Fig. 7). For instance, Mucin-5AC~~,~~ and Mucin-5B ~~and~~ (i.e. the major ~~component~~components of jelly core proteins), were relatively highly expressed in the [T] individuals. However~~, in general~~, expression of the putative jelly coat genes was very diverse across the six individuals in our data set (Fig. ~~7).~~ 7). More detailed analyses on coding versus regulatory variation underlying this heterogeneity as well as SNP genotyping of allelic variation across the acidification gradient are important future avenues. However, it is important to note that the current data only provides the first step that in future allows to test to what extent variation is due to coding and/or transcriptional differences in the maternally influenced jelly coat (which, in turn, affects embryonic survival).

The potential complexity of the jelly phenotype and its function is also highlighted in the high degree of variation in expression of the putative jelly coat genes across different individuals in our study (Fig. 7). Given the many important roles that the egg jelly coats play in sperm-egg interactions, pathogen defense and responses to various environmental stressors (reviewed in ~~(Menkhorst & Selwood 2008; Shu et al. 2015b)~~(Menkhorst & Selwood 2008; Shu et al. 2015b)), this complexity is not surprising. Follow up work establishing the link between the genotype (jelly coat coding genes and variation in the expression) – phenotype (jelly coat glycome) - fitness (embryonic acid tolerance) map are needed to confirm the role of the candidate genes is ongoing in our lab.

~~Conclusion~~

Conclusions

In conclusion, we identified several mucin and O-linked glycosylation genes that are highly expressed in the oviduct of *R. arvalis* – and show high heterogeneity in expression. Given the role of *R. arvalis* in a broad range of evolutionary ecological questions, we believe this transcriptomic dataset ~~and~~together with the predicted SSR and SNP markers (Supporting information) will benefit future studies of molecular ecology and evolution in natural populations. We further hope that our oviduct transcriptomes ~~lays~~lay the ground for future studies on the molecular evolution of jelly coat genes, thereby contributing to the emerging field of glycobiology in evolutionary ecology ~~(Shu et al. 2015b; Springer & Gagneux 2016)~~(Shu et al. 2015b; Springer & Gagneux 2016), as well as studies on how these genes contribute to responses to natural selection at early life stages and adaptive divergence of local populations in particular. In particular, we hope that the genes identified here will aid in disentangling the genetic architecture of egg coat evolution and adaptive maternal effects.

~~Declarations~~

Acknowledgements

Formatted: Font: 14 pt

We thank Anssi Laurila for generous access to research facilities and Beatrice Lindgren for invaluable help in the field and ~~Dr.~~laboratory. We thank particularly Ellen Menkhorst, Daniel Jeffries, Alice Dennis ~~for~~and an anonymous reviewer for valuable comments on earlier drafts of this manuscript ~~and~~as well as the Genetic Diversity Center (GDC) of ETH-Zurich for advice on genetic analyses. The experiments were conducted under permissions from the Västra Götaland county board and the Ethical committee for animal experiments in Uppsala County. ~~This study was supported by Swiss National Science foundation (to KR).~~Longfei Shu is currently a Simons Foundation Fellow of the Life Sciences Research Foundation.

~~Data Accessibility~~

~~Transcriptome sequences will be deposited in the NCBI Sequence Read Archive (SRA) database.~~

~~Author Contributions~~

~~LS and KR conceived and planned the study. LS performed the experiments. Both authors discussed the results and commented on the manuscript. Both authors read and approved the final manuscript.~~

~~Competing interests~~

~~The authors declare that they have no competing interests.~~

References

Bernardo J. 1996. The particular maternal effect of propagule size, especially egg size: Patterns, models, quality of evidence and interpretations. *American Zoologist* 36:216-236.

Birol I, Behsaz B, Hammond SA, Kucuk E, Veldhoen N, and Helbing CC. 2015. De novo Transcriptome Assemblies of Rana (Lithobates) catesbeiana and Xenopus laevis Tadpole Livers for Comparative Genomics without Reference Genomes. *Plos One* 10. ARTN e0130720 10.1371/journal.pone.0130720

~~Bolek MG, and Janovy J. 2007. Evolutionary avenues for, and constraints on, the transmission of frog lung flukes (Haematoloechus spp.) in dragonfly second intermediate hosts. *Journal of Parasitology* 93:593-607. Doi 10.1645/Ge-1011r.1~~

Brelsford A, Lavanchy G, Sermier R, Rausch A, and Perrin N. 2017. Identifying homomorphic sex chromosomes from wild-caught adults with limited genomic resources. *Molecular Ecology Resources* 17:752-759. 10.1111/1755-0998.12624

Calboli FCF, Fisher MC, Garner TWJ, and Jehle R. 2011. The need for jumpstarting amphibian genome projects. *Trends in Ecology & Evolution* 26:378-379. 10.1016/j.tree.2011.04.006

Claw KG, and Swanson WJ. 2012. Evolution of the Egg: New Findings and Challenges. In: Chakravarti A, and Green E, eds. *Annual Review of Genomics and Human Genetics, Vol 13*. Palo Alto: Annual Reviews, 109-125.

Coppin A, Maes E, Flahaut C, Coddeville B, and Strecker G. 1999. Acquisition of species-specific O-linked carbohydrate chains from oviducal mucins in Rana arvalis - A case study. *European Journal of Biochemistry* 266:370-382.

Delplace F, Maes E, Lemoine J, and Strecker G. 2002. Species specificity of O-linked carbohydrate chains of the oviducal mucins in amphibians: structural analysis of neutral oligosaccharide alditols released by reductive beta-elimination from the egg-jelly coats of Rana clamitans. *Biochemical Journal* 363:457-471.

Egea-Serrano A, Hangartner S, Laurila A, and Rasanen K. 2014. Multifarious selection through environmental change: acidity and predator-mediated adaptive divergence in the moor frog (Rana arvalis). *Proceedings of the Royal Society B-Biological Sciences* 281. Unsp 20133266 Doi 10.1098/Rspb.2013.3266

Elewa A, Wang H, Talavera-Lopez C, Joven A, Brito G, Kumar A, Hameed LS, Penrad-Mobayed M, Yao ZY, Zamani N, Abbas Y, Abdullayev I, Sandberg R, Grabherr M, Andersson B, and Simon A. 2017. Reading and editing the Pleurodeles waltl genome reveals novel features of tetrapod regeneration. *Nature Communications* 8. Artn 2286 10.1038/S41467-017-01964-9

Gomez-Mestre I, Touchon JC, and Warkentin KM. 2006. Amphibian embryo and parental defenses and a larval predator reduce egg mortality from water mold. *Ecology* 87:2570-2581. Doi 10.1890/0012-9658(2006)87[2570:Aeapda]2.0.Co;2

Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, and Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* 36:3420-3435. 10.1093/nar/gkn176

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, and Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644-U130. 10.1038/nbt.1883

Guerardel Y, Kol O, Maes E, Lefebvre T, Boilly B, Davril M, and Strecker G. 2000. O-glycan variability of egg-jelly mucins from Xenopus laevis: characterization of four phenotypes that differ by the terminal glycosylation of their mucins. *Biochemical Journal* 352:449-463.

Hammond SA, Warren RL, Vandervalk BP, Kucuk E, Khan H, Gibb EA, Pandoh P, Kirk H, Zhao YJ, Jones M, Mungall AJ, Coope R, Pleasance S, Moore RA, Holt RA, Round JM, Ohora S, Walle BV, Veldhoen N, Helbing CC, and Birol I. 2017. The North American bullfrog draft genome provides insight into hormonal regulation of long noncoding RNA. *Nature Communications* 8. Artn 1433 10.1038/S41467-017-01316-7

Hangartner S, Laurila A, and Rasanen K. 2011. Adaptive divergence of the moor frog (Rana arvalis) along an acidification gradient. *Bmc Evolutionary Biology* 11. 366 10.1186/1471-2148-11-366

Hangartner S, Laurila A, and Rasanen K. 2012. The quantitative genetic basis of adaptive divergence in the moor frog (Rana arvalis) and its implications for gene flow. *Journal of Evolutionary Biology*~~.~~ 25:1587-1599. 10.1111/j.1420-9101.2012.02546.x

Hedrick JL. 2008. Anuran and pig egg zona pellucida glycoproteins in fertilization and early development. *International Journal of Developmental Biology* 52:683-701. DOI 10.1387/ijdb.082580jh

Hedrick JL, and Nishihara T. 1991. Structure and Function of the Extracellular-Matrix of Anuran Eggs. *Journal of Electron Microscopy Technique* 17:319-335.

Helbing CC. 2012. The metamorphosis of amphibian toxicogenomics. *Front Genet* 3:37. 10.3389/fgene.2012.00037

Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu SQ, Taher L, Blitz IL, Blumberg B, Dichmann DS, Dubchak I, Amaya E, Detter JC, Fletcher R, Gerhard DS, Goodstein D, Graves T, Grigoriev IV, Grimwood J, Kawashima T, Lindquist E, Lucas SM, Mead PE, Mitros T, Ogino H, Ohta Y, Poliakov AV, Pollet N, Robert J, Salamov A, Sater AK, Schmutz J, Terry A, Vize PD, Warren WC, Wells D, Wills A, Wilson RK, Zimmerman LB, Zorn AM, Grainger R, Grammer T, Khokha MK, Richardson PM, and Rokhsar DS. 2010. The Genome of the Western Clawed Frog Xenopus tropicalis. *Science* 328:633-636. 10.1126/science.1183670

Hof C, Araujo MB, Jetz W, and Rahbek C. 2011. Additive threats from pathogens, climate and land-use change for global amphibian diversity. *Nature* 480:516-~~U137~~519. 10.1038/nature10650

Houle D, Govindaraju DR, and Omholt S. 2010. Phenomics: the next challenge. *Nature Reviews Genetics* 11:855-866. 10.1038/nrg2897

Iseli C, Jongeneel CV, and Bucher P. 1999. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc Int Conf Intell Syst Mol Biol*:138-148.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, and Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Research* 36:~~D480~~480-484. 10.1093/nar/gkm882

Lang TG, Klasson S, Larsson E, Johansson MEV, Hansson GC, and Samuelsson T. 2016. Searching the Evolutionary Origin of Epithelial Mucus Protein Components-Mucins and FCGBP. *Molecular Biology and Evolution* 33:1921-1936. 10.1093/molbev/msw066

Li RQ, Li YR, Kristiansen K, and Wang J. 2008. SOAP: short oligonucleotide alignment program. *Bioinformatics* 24:713-714. 10.1093/bioinformatics/btn025

Lopez-Maestre H, Brinza L, Marchet C, Kielbassa J, Bastien S, Boutigny M, Monnin D, El Filali A, Carareto CM, Vieira C, Picard F, Kremer N, Vavre F, Sagot MF, and Lacroix V. 2016. SNP calling from RNA-seq data without a reference genome: identification, quantification, differential analysis and impact on the protein sequence. *Nucleic Acids Research* 44:e148. 10.1093/nar/gkw655

Menkhorst E, and Selwood L. 2008. Vertebrate Extracellular Preovulatory and Postovulatory Egg Coats. *Biology of Reproduction* 79:790-797. 10.1095/biolreprod.108.068551

Mitchell-Olds T, Willis JH, and Goldstein DB. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* 8:845-856. Doi 10.1038/Nrg2207

Morin PA, Luikart G, Wayne RK, and Grp SW. 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* 19:208-216. 10.1016/j.tree.2004.01.009

Mortazavi A, Williams BA, McCue K, Schaeffer L, and Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621-628. 10.1038/nmeth.1226

Mousseau TA, and Fox CW. 1998. The adaptive significance of maternal effects. *Trends in Ecology & Evolution* 13:403-407.

Nadeau NJ, and Jiggins CD. 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics* 26:484-492. DOI 10.1016/j.tig.2010.08.004

Nowoshilow S, Schloissnig S, Fei JF, Dahl A, Pang AWC, Pippel M, Winkler S, Hastie AR, Young G, Roscito JG, Falcon F, Knapp D, Powell S, Cruz A, Cao H, Habermann B, Hiller M, Tanaka EM, and Myers EW. 2018. The axolotl genome and the evolution of key tissue formation regulators. *Nature* 554:50-55. 10.1038/nature25458

Palumbi SR. 2009. Speciation and the evolution of gamete recognition genes: pattern and process. *Heredity* 102:66-76. Doi 10.1038/Hdy.2008.104

Papantonis A, Swevers L, and Iatrou K. 2015. Chorion Genes: A Landscape of Their Evolution, Structure, and Regulation. *Annual Review of Entomology* 60:177-194. 10.1146/annurev-ento-010814-020810

Persson M, Räsänen K, Laurila A, and Merilä J. 2007. Maternally determined adaptation to acidity in Rana arvalis: Are laboratory and field estimates of embryonic stress tolerance congruent? *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 85:832-838. Doi 10.1139/Z07-064

Price SJ, Garner TWJ, Balloux F, Ruis C, Paszkiewicz KH, Moore K, and Griffiths AGF. 2015. A de novo Assembly of the Common Frog (Rana temporaria) Transcriptome and Comparison of Transcription Following Exposure to Ranavirus and Batrachochytrium dendrobatidis. *Plos One* 10. ARTN e0130500 10.1371/journal.pone.0130500

Qiao L, Yang WZ, Fu JZ, and Song ZB. 2013. Transcriptome Profile of the Green Odorous Frog (Odorrana margaretae). *Plos One* 8. ARTN e75211 10.1371/journal.pone.0075211

Räsänen K, and Kruuk LEB. 2007. Maternal effects and evolution at ecological time-scales. *Functional Ecology* 21:408-421. DOI 10.1111/j.1365-2435.2007.01246.x

Räsänen K, Laurila A, and Merilä J. ~~2003.~~2003a. Geographic variation in acid stress tolerance of the moor frog, Rana arvalis. I. Local adaptation. *Evolution* 57:352-362.

Räsänen K, Laurila A, and Merilä J. 2003b. Geographic variation in acid stress tolerance of the moor frog, Rana arvalis. II. Adaptive maternal effects. *Evolution* 57:363-371.

Richter-Boix A, Quintela M, Segelbacher G, and Laurila A. 2011. Genetic analysis of differentiation among breeding ponds reveals a candidate gene for local adaptation in Rana arvalis. *Molecular Ecology* 20:1582-1600. 10.1111/j.1365-294X.2011.05025.x

Robertson LS, and Cornman RS. 2014. Transcriptome resources for the frogs Lithobates clamitans and Pseudacris regilla, emphasizing antimicrobial peptides and conserved loci for phylogenetics. *Molecular Ecology Resources* 14:178-183. 10.1111/1755-0998.12164

Santos JC, Tarvin RD, O'Connell LA, Blackburn DC, and Coloma LA. 2018. Diversity within diversity: Parasite species richness in poison frogs assessed by transcriptomics. *Molecular Phylogenetics and Evolution* 125:40-50. https://doi.org/10.1016/j.ympev.2018.03.015

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, and Ragg T. 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *Bmc Molecular Biology* 7. Artn 3 10.1186/1471-2199-7-3

Schunter C, Garza JC, Macpherson E, and Pascual M. 2014. SNP development from RNA-seq data in a nonmodel fish: how many individuals are needed for accurate allele frequency prediction? *Molecular Ecology Resources* 14:157-165. Doi 10.1111/1755-0998.12155

Sears BF, Schlunk AD, and Rohr JR. 2012. Do Parasitic Trematode Cercariae Demonstrate a Preference for Susceptible Host Species? *Plos One* 7. ARTN e51012 10.1371/journal.pone.0051012

Shu L, Laurila A, and Rasanen K. 2015a. Acid stress mediated adaptive divergence in ion channel function during embryogenesis in Rana arvalis. *Scientific Reports* 5:14201. 10.1038/srep14201

Shu L, Laurila A, Suter MJF, and Räsänen K. 2016. Molecular phenotyping of maternally mediated parallel adaptive divergence within Rana arvalis and Rana temporaria. *Molecular Ecology*~~.~~ 25:4564-4579. 10.1111/mec.13786

Shu L, Suter M, and Rasanen K. 2015b. Evolution of egg coats: linking molecular biology and ecology. *Molecular Ecology* 24:4052-4073. 10.1111/mec.13283

Shu L, Suter MF, Laurila A, and Räsänen K. 2015c. Mechanistic basis of adaptive maternal effects: egg jelly water balance mediates embryonic adaptation to acidity in Rana arvalis. *Oecologia*~~:1-12~~ 179:617-628. 10.1007/s00442-015-3332-4

Springer SA, and Gagneux P. 2016. Glycomics: revealing the dynamic ecology and evolution of sugar molecules. *J Proteomics* 135:90-100. 10.1016/j.jprot.2015.11.022

Strecker G, Coppin A, Florea D, Maes E, and Cogalniceanu D. 2003. Comparative study of carbohydrate chains released from the oviducal mucins of the two very closely related amphibian species Bombina bombina and Bombina variegata. *Biochimie* 85:53-64. Doi 10.1016/S0300-9084(03)00021-X

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, and Waller RW. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783-1786. 10.1126/science.1103538

Sun YB, Xiong ZJ, Xiang XY, Liu SP, Zhou WW, Tu XL, Zhong L, Wang L, Wu DD, Zhang BL, Zhu CL, Yang MM, Chen HM, Li F, Zhou L, Feng SH, Huang C, Zhang GJ, Irwin D, Hillis DM, Murphy RW, Yang HM, Che J, Wang J, and Zhang YP. 2015. Whole-genome sequence of the Tibetan frog Nanorana parkeri and the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of Sciences of the United States of America* 112:~~E1257-E1262~~1257-1262. 10.1073/pnas.1501764112

Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, and Conesa A. 2011. Differential expression in RNA-seq: A matter of depth. *Genome Research* 21:2213-2223. 10.1101/gr.124321.111

Todd EV, Black MA, and Gemmell NJ. 2016. The power and promise of RNA-seq in ecology and evolution. *Molecular Ecology* 25:1224-1241. 10.1111/mec.13526

Tong ZB, Gold L, Pfeifer KE, Dorward H, Lee E, Bondy CA, Dean J, and Nelson LM. 2000. Mater, a maternal effect gene required for early embryonic development in mice. *Nature Genetics* 26:267-268.

Van Buskirk J, and Arioli M. 2005. Habitat specialization and adaptive phenotypic divergence of anuran populations. *Journal of Evolutionary Biology* 18:596-608. 10.1111/j.1420-9101.2004.00869.x

Wang Z, Gerstein M, and Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10:57-63. Doi 10.1038/Nrg2484

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, and Wade MJ. 1998. Evolutionary consequences of indirect genetic effects. *Trends in Ecology & Evolution* 13:64-69. Doi 10.1016/S0169-5347(97)01233-0

Wong JL, and Wessel GM. 2006. Defending the zygote: Search for the ancestral animal block to polyspermy. *Current Topics in Developmental Biology* 72:1~~-+.~~-151. Doi 10.1016/S0070-2153(05)72001-9

Wu XM, Viveiros MM, Eppig JJ, Bai YC, Fitzpatrick SL, and Matzuk MM. 2003. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nature Genetics* 33:187-191. 10.1038/ng1079

Zhang M, Li YT, Yao BJ, Sun MY, Wang ZW, and Zhao Y. 2013. Transcriptome Sequencing and de novo Analysis for Oviductus Ranae of Rana chensinensis Using Illumina RNA-Seq Technology. *Journal of Genetics and Genomics* 40:137-140. 10.1016/j.jgg.2013.01.004