Candidate genes involved in the evolution of viviparity: a RAD sequencing experiment in the lizard *Zootoca vivipara* (Squamata: Lacertidae)

LUCA CORNETTI[1,2,3], OLIVER W. GRIFFITH[4], ANDREA BENAZZO[2], ALEX PANZIERA[2], CAMILLA M. WHITTINGTON[5,6], MICHAEL B. THOMPSON[5], CRISTIANO VERNESI[1] and GIORGIO BERTORELLE[2*]

[1]*Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy*
[2]*Department of Life Sciences and Biotechnology, University of Ferrara, via Borsari 46, 44121 Ferrara, Italy*
[3]*Universität Basel, Zoologisches Institut, Evolutionsbiologie, Vesalgasse 1, 4051 Basel, Switzerland*
[4]*Department of Ecology and Evolutionary Biology, Yale University, 850 West Campus Drive, West Haven, CT, 06516, USA*
[5]*School of Life and Environmental Sciences, University of Sydney, Heydon-Laurence Building (A08), Sydney, NSW 2006, Australia*
[6]*Sydney School of Veterinary Science, University of Sydney, Sydney, Regimental Dr, Camperdown NSW 2050, Australia*

Viviparity has evolved from oviparity at least 150 independent times in vertebrates. More than 80% of these transitions have occurred in squamate reptiles, where both reproductive modes are rarely seen in different populations of the same species. This condition (bimodal reproduction) is ideal for studying the physiological and morphological changes underpinning the evolution of reproductive mode, and their genetic determinants. Here we analysed the genomes of *Zootoca vivipara* populations with either oviparous or viviparous reproduction using a RAD sequencing approach. No signature of interbreeding between oviparous and viviparous individuals was found. We conservatively identified 22 annotated coding sequences in genes potentially associated with parity mode differences. Six of these genes are transcription regulators that are also expressed in reproductive tissues of mammals and reptiles, suggesting that changes in gene expression are important for the evolution of viviparity. Using a more inclusive approach based on contigs mapping in either coding or non-coding regions, 45 genes were identified. Twelve of these candidate genes are transcription regulators and four encode protease enzymes. We propose that the evolution of proteases may support morphological changes to the uterus during pregnancy. This study provides the foundation for further experimental studies of the genetic basis of parity mode in *Z. vivipara*.

ADDITIONAL KEYWORDS: evolutionary genomics – gene–phenotype association – oviparity – population genetics – reproductive strategies – viviparity.

INTRODUCTION

Reproductive mode fundamentally affects life-history patterns of living organisms, and it is associated with the evolution of a complex assortment of morphological structures and physiological functions. Scientists have always been fascinated by the variation in reproductive mode across kingdoms (Angelini & Ghiara, 1984; Holsinger, 2000; Touchon & Warkentin, 2008), and even within the same species (Guillette, 1982; Adams *et al.*, 2007; Sandrock, Schirrmeister & Vorburger, 2011) but the understanding of the suite of genomic changes associated with, and possibly driving, this transition is still very limited.

*Corresponding author. E-mail: ggb@unife.it

Among vertebrates, the two main types of reproductive mode are oviparity (mothers lay eggs) and viviparity (embryos develop inside the body of the parent, which then gives birth to live offspring). More than 150 independent transitions to viviparity from an oviparous ancestor have been described (Blackburn, 2015; Griffith *et al.*, 2015), with physical changes required to allow for pregnancy, including the development of mechanisms to exchange respiratory gasses, water and nutrients, and to regulate immunity (Thompson & Speake, 2006; Van Dyke, Brandley & Thompson, 2014). These changes can be achieved by changes in gene regulation, which alter the physical properties of cells in reproductive tissues, or by re-patterning of reproductive tissues so that they behave differently (Griffith *et al.*, 2016b; Griffith & Wagner, 2017). An example of the latter is the reduction in uterine glands in viviparous *Zootoca vivipara*, which resulted in reduced shell thickness and allowed greater exchange of materials through pregnancy (Heulin *et al.*, 2005). By comparing specific traits, genes and genomes between closely related viviparous and oviparous taxa, it is possible to identify the mechanisms that underlie the evolutionary transition from oviparity to viviparity (Murphy & Thompson, 2011).

Within Squamata (lizards and snakes), there are also peculiar cases of closely related lineages that display different reproductive modes (i.e. oviparous and viviparous populations within the same species). In fact, four species show geographic variation in reproductive mode, namely the Australian scincid lizards *Lerista bougainvillii* and *Saiphos equalis* (Smith & Shine, 1997; Qualls & Shine, 2006), the South American water snake *Helicops angulatus* (Braz, Scartozzoni & Almeida-Santos, 2016) and the Eurasian lacertid lizard *Z. vivipara* (Surget-Groba *et al.*, 2006). These species provide ideal models for studying morphological and physiological modifications as well as the genetic processes underlying the transition from oviparity to viviparity. The focus of the present study is *Z. vivipara*.

Zootoca vivipara is distributed throughout Europe and Asia, and it is one of the reptile species with the widest distribution. The vast majority of populations, from Japan to Central Western Europe, including the British Isles and Scandinavia, are viviparous and are classified as *Zootoca vivipara vivipara*. Distinct but related viviparous mitochondrial clades are classified as *Zootoca vivipara pannonica* and *Zootoca vivipara sachalinensis* (Surget-Groba *et al.*, 2001), but in this paper, we use only the nominal subspecies *Z. v. vivipara* to identify the viviparous group. Two disjunct oviparous populations exist and are classified as *Zootoca vivipara louislantzi* and *Zootoca vivipara carniolica*. *Zootoca vivipara louislantzi* occurs in the Pyrenees, and its overall genetic affinity within the viviparous clade supports the hypothesis of a secondary reversion

to the oviparous reproductive mode (Surget-Groba *et al.*, 2006). *Zootoca vivipara carniolica* is distributed in Alpine regions in northern Italy, southern Austria, Slovenia and Croatia and is considered the ancestral oviparous form (Surget-Groba *et al.*, 2001). The alternative hypothesis of ancestrality of both oviparous groups, accompanied by multiple independent origins of viviparity, although less parsimonious, cannot be excluded, and more developmental and physiological evidence is required to separate these hypotheses (Griffith *et al.*, 2015).

An important aspect of the phylogeography within this species is that viviparous lineages also occur beyond the Arctic Circle, whereas oviparous lineages are present only in the southern margin of species distribution (Sillero *et al.*, 2014), where average temperatures are warmer. This scenario is in agreement with the 'cold-climate hypothesis' that posits that evolution of viviparity is more likely at low environmental temperatures, and also with the observation that live-bearing species are generally distributed in colder habitats than egg-laying taxa (Shine, 2005; Rodríguez-Díaz & Braña, 2012; Cornetti *et al.*, 2015a). In *Z. vivipara*, viviparity probably evolved as a consequence of selective pressure caused by cold climatic conditions during Pleistocene glacial phases (Surget-Groba *et al.*, 2001), and the newly evolved reproductive mode then allowed viviparous populations to colonize the vast majority of Eurasia (Cornetti *et al.*, 2014).

The evolution of viviparity requires specific developmental changes to support pregnancy. These changes include regulatory mechanisms that maintain embryos *in utero* through development, mechanisms to support embryonic gas exchange through pregnancy and mechanisms to transport calcium to embryos in the absence of eggshell-derived calcium (Thompson & Speake, 2006; Stewart, Ecay & Heulin, 2009; Murphy & Thompson, 2011; Stewart *et al.*, 2011; Stewart, 2013). The evolution of innovations in organisms is achieved through both changes in the peptide sequence of genes, resulting in proteins with novel functions, and changes in gene expression (Rawn & Cross, 2008). By inducing the expression of genes normally expressed elsewhere in the organism, tissues can take on novel functions (True & Carroll, 2002). Comparative studies between oviparous and viviparous populations of *Z. vivipara* failed to identify differences in the expression of interleukin genes and their receptors (Paulesu *et al.*, 2005). Furthermore, studies on hormone-related genes and angiogenic factors have also failed to identify expression profiles correlated with reproductive mode (Whittington *et al.*, 2015a; Griffith *et al.*, 2017). However, transcriptome sequencing of the uterus of the African ocellated skink, *Chalcides ocellatus*, and the Australian Southern grass skink, *Pseudemoia*

entrecasteauxii, reveals large changes in gene regulation through pregnancy (Brandley *et al.*, 2012; Griffith *et al.*, 2016b). Given that viviparity requires substantial changes in gene expression to support the function of pregnancy, a plausible hypothesis to test is that genetic changes responsible for the evolution of viviparity mostly occur in regions of the genome responsible for gene regulation.

Here we used RAD (restriction-site associated DNA) sequencing of *Z. vivipara* from populations showing different reproductive modes. We initially compared the new genomic data with previous mtDNA and microsatellite inference to better understand the geographic structure of this species. We subsequently focused on the identification of genomic regions or genes possibly associated with the different reproductive modes, by contrasting allele frequencies of oviparous and viviparous populations and conducting genotype–phenotype association tests.

MATERIAL AND METHODS

SAMPLING, DNA EXTRACTION AND LIBRARY PREPARATION

Forty tail tips of *Z. vivipara* adults were collected to cover most of the mitochondrial lineages described in Surget-Groba *et al.* (2006). The capture and handling of lizards complied with national and international ethical guidelines, and the Italian Ministry of Environment and the Environmental Unit of the Autonomous Province of Trento approved capture, handling and tissue sampling (DPN/2D/2003/2267 and 4940–57/B-09-U265-LS-fd). We used tail tissues from individuals already analysed for phylogeographic studies (Surget-Groba *et al.*, 2006; Cornetti *et al.*, 2014, 2015a). All the individuals were previously sequenced at mitochondrial DNA gene *cytochrome b* for subspecies assignment because no single morphometric trait can distinguish the *Z. vivipara* subspecies (Guillaume *et al.*, 2006). Our sample set included ten individuals of the oviparous subspecies *Z. v. carniolica* from the European Alps (mtDNA: clade A), seven individuals of the oviparous subspecies *Z. v. louislantzi* from the Pyrenees (mtDNA: cladeB), 17 individuals of the viviparous subspecies *Z. v. vivipara* from the European Alps (mtDNA: clade E), plus six additional viviparous individuals of *Z. v. vivipara* representing two additional mtDNA clades (four from clade D and two from clade F). Most of the described mtDNA clades are thus considered (see Supporting Information, Table S1).

Genomic DNA was extracted using the QIAGEN DNeasy Blood and Tissue Kit (QIAGEN Inc., Hilden, Germany). DNA was treated with RNaseA (QIAGEN) and quantified with the fluorometer Qubit 2.0 (Invitrogen). We then prepared libraries for reducing the complexity of the genome by fragmenting genomic DNA with a restriction enzyme (RAD sequencing [Baird *et al.*, 2008]). We digested, with SbfI, 1 µg of DNA for each individual sample in a 50 µL reaction volume. P1 adapter, containing unique barcode, was ligated onto complementary compatible ends for each sample. Individually barcoded samples were pooled and then sheared to an average size of 500 bp using the ultrasonicator Covaris S220 (Covaris Inc., Woburn, MA, USA). A size selection by means of agarose gel extraction was performed to restrict the size range of fragments to that which can be efficiently sequenced on an Illumina flow cell (300–500 bp). Library preparation was completed after ligating P2 adapters that allowed amplification of fragments that incorporated both P1 and P2. The library was run on one lane using Illumina HiSeq2000 and the 100 bp paired-end protocol, at the GenePool (Edinburgh, Scotland).

SINGLE-NUCLEOTIDE POLYMORPHISM GENOTYPING

Quality of the raw Illumina reads was examined with FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), especially for evaluating the possible decrease in quality score along the read length. In fact, reads were trimmed after 75 bp to avoid including low-quality sequence fragments; barcodes and restriction sites were also trimmed resulting in a final read length of 64 bp. Reads with ambiguous barcodes, ambiguous restriction sites or showing low quality were discarded. The quality of the reads was analysed using sliding windows of a length corresponding to 15% of the length of the read. The average quality score was calculated within each window. Whenever the base call accuracy dropped below 90%, corresponding to a Phred score of 10, the read was discarded. Additionally, to have a non-redundant data set, reads identified as PCR clones (i.e. identical in both paired-ends) were reduced to a single copy. The remaining single-end reads (sequences flanking the restriction sites) were then used for single-nucleotide polymorphism (SNP) calling with the software Stacks v 1.02 (Catchen *et al.*, 2013). More specifically, the Perl script denovo_map.pl was used (1) to align single-end reads into exactly matching stacks, setting six reads as the minimum coverage for each stack; (2) to compare stacks for obtaining a set of loci and call SNPs using a maximum likelihood framework (Hohenlohe *et al.*, 2010); and (3) to build a catalogue of loci against which all samples were compared. The programme *populations* implemented in Stacks was used to export genotypes.

Genomic divergence among individuals and groups

A triangular matrix of genetic distances computed as 1 minus the fraction of shared alleles among pairs of individuals was visualized by a multidimensional scaling (MDS) using *cmdscale* and by a neighbor-joining (NJ) tree using *ape* (R Core Development Team, 2015). The genomic composition of each individual sampled in the Alps (where the oviparous and the viviparous forms are geographically close and overlap in some areas) was analysed using the methods implemented in STRUCTURE (Pritchard, Stephens & Donnelly, 2000) and Treemix (Pickrell & Pritchard, 2012). The number of groups in Structure (K) was allowed to vary from 2 to 4, running ten independent runs (for each K) consisting of 500 000 iterations after a burn-in period of 100 000. The analyses were performed twice, allowing a maximum of either 20 or 30% of missing data. Treemix was run allowing zero to five migration edges among the major five groups and computing the fraction of explained variance for each model (a model being a population divergence tree with different numbers of migration edges).

Defining SNPs possibly involved in the switch in reproductive mode

We defined SNPs that were identified by at least two of three different methods as candidate loci potentially associated with transitions between reproductive modes. We used in this analysis all the SNPs with up to 50% of missing data.

The first method was developed specifically for this species, where the transition between oviparity and viviparity occurred more than once due to a reversal event in *Z. v. louislantzi* from a viviparous ancestor (Surget-Groba *et al.*, 2006) or, less likely, to multiple origins of viviparity. The method controls for the global evolutionary divergence between oviparous and viviparous subspecies not related to the different reproductive modes, by plausibly assuming that the same relevant loci changed during the reproductive mode transitions occurred within this species. Candidate SNPs were conservatively identified when they simultaneously satisfied the following extreme conditions: $F_{st} \geq 0.5$ between *Z. v. vivipara* (V) and *Z. v. carniolica* (O); $F_{st} \geq 0.5$ between *Z. v. vivipara* (V) and *Z. v. louislantzi* (O); and $F_{st} \leq 0.05$ between *Z. v. carniolica* (O) and *Z. v. louislantzi* (O). The F_{st} distributions in the three comparisons, with the result produced by the chosen cutoff values, are reported in Supporting Information, Figure S1.

Second, we used the method genome-wide efficient mixed-model association (GEMMA) (Zhou & Stephens, 2012), which calculates statistical genotype–phenotype association implementing the GEMMA algorithm. The univariate linear mixed model with Wald test was used, providing as input files the subset of markers with less than 50% of missing data and a centred genotype relatedness matrix estimated from our genotypes. P-values from the Wald test were corrected for multiple testing (Benjamini & Hochberg, 1995), and SNPs with corrected P-values lower than 0.05 were retained as possibly associated with the reproductive trait.

Third, we tested for significant association between genotypes and the reproductive mode with TASSEL 4.0 (Bradbury *et al.*, 2007). We initially estimated the genetic composition of each individual and the most probable number K of genetically homogeneous groups using STRUCTURE v2.3.4 (Pritchard *et al.*, 2000) running ten independent runs consisting of 500 000 iterations after a burn-in period of 100 000 for all K values between 1 and 5. This analysis was conducted on a subset of 3000 randomly selected unlinked SNPs among the total number of polymorphisms obtained (see below). The most likely number of K ($K = 3$) was estimated using STRUCTURE HARVESTER (Evanno, Regnaut & Goudet, 2005; Earl & VonHoldt, 2012). The resultant Q-matrix and a kinship matrix were used to correct for population structure in the mixed linear model (Zhang *et al.*, 2010) applied for genotype–phenotype association implemented in TASSEL. The genotype association with the trait (reproductive mode, oviparity or viviparity) was considered to be significant if $P < 0.05$, after correction for multiple testing (Benjamini & Hochberg, 1995).

Linking polymorphisms under selection with annotated genes

Both single-end RAD-seq loci and the mini-contigs generated using the paired-end reads were blasted against a DNA data set of 21 913 *Anolis carolinensis* complete genes [exons and introns, extracted from ENSEMBL database using BIOMART (Smedley *et al.*, 2015)] with dc-megablast (Morgulis *et al.*, 2008), to identify putative homologous genes between the two species. Mini-contigs for each RAD locus were assembled from paired-end reads using the Perl script exec_velvet.pl implemented in Stacks. Only sequences showing at least 80% of identity for at least 40% of their length were retained. We also followed a protein-based reciprocal blast approach to identify putative homologous genes. All the possible translations of the loci putatively under selection were blasted with tBLASTx (Camacho *et al.*, 2009) to the translated *A. carolinensis* transcriptome. The best hits were blasted back (again with tBLASTx) to all the translated RAD contigs. Two lists of putatively homologous genes between *Z. vivipara* and *A. carolinensis* were then obtained. In the c-list (conservative list), we

included only the genes showing an *E*-value < 10^{-6} in the forward or the reverse fragment under the DNA and the protein approaches. We included in the i-list (inclusive list) the genes identified by either the DNA or the protein approaches (i.e. the genes in the c-list plus the genes identified only by one approach). PantherDB (Thomas *et al.*, 2003) was used to identify the Gene Ontology (GO) category of each gene in the p-list. A test of overrepresentation was finally performed on the p-list and the i-list using an integrated tool in PantherDB. Specifically, we compared the proportions of GO terms in the set of loci putatively associated with the transition in reproductive mode with the proportions observed in the *A. carolinensis* genome.

RESULTS

SNP GENOTYPING

The RAD sequencing experiment produced 146 439 826 paired-end raw reads. We discarded reads with ambiguous barcodes (24 270), ambiguous restriction sites (31 474 031) and reads with low base call accuracy (5 830 599). PCR clones, corresponding to 59.4% of the total, were reduced to a single copy. The accession for the Illumina reads in SRA is SRP109076, while the BioProject accession is PRJNA390236. Stack files will be available upon request. Genotypic calling was performed using the retained 22 197 925 single-end reads. According to the decay in quality score along read length indicated by FastQC, we trimmed the last 25 bp to guarantee reliable SNP calling. We identified a total of 45 151 contigs and 80 792 SNPs (minimum coverage of 6× in at least two individuals), selected from the single-end reads showing no more than five polymorphisms. These SNPs were used to describe the overall genetic variation within and between lineages. Details of the reads and the average coverage per individual are given in Supporting Information, Table S1.

PHYLOGEOGRAPHIC ANALYSIS

The MDS plot (Fig. 1A) supports the existence of three major genomic groups, clearly corresponding to the different reproductive modes and the different evolutionary history of the oviparous populations that has been suggested for this species (Surget-Groba *et al.*, 2001). Along the first axis, the oviparous *Z. v. carniolica* (the group with the ancestral oviparous state) is separated from a group composed by the viviparous *Z. v. vivipara* individuals and the oviparous *Z. v. louislantzi*, with the latter probably having a secondary derived mode of reproduction evolved from a viviparous ancestor (Surget-Groba *et al.*, 2006). The individuals sampled in the same Eastern Alps area but with different reproductive mode (*Z. v. carniolica* and *Z. v. vivipara*, see Fig. 1A) are genetically very different, supporting the hypothesis that hybridization does not occur (or does not produce offspring) in natural sympatric conditions (Cornetti *et al.*, 2015b). *Zootoca v. vivipara* and *Z. v. louislantzi* then form two distinct groups separated along the *Y*-axis (Fig. 1A). There is also genetic substructure, especially in the individual tree (Fig. 1B), in *Z. v. vivipara*, where individuals belonging to different mtDNA clades and sampled in different countries show some level of genetic heterogeneity. The total number of SNPs used in the MDS analysis and in the NJ tree was 80 792, and the average number of SNPs used in the pairwise comparisons was 7795.0. The individuals sampled in the Alps and analysed with STRUCTURE (ten belonging to the *Z. v. carniolica* mtDNA oviparous lineage, and 19 to the *Z. v. vivipara* mtDNA viviparous lineage) can be unequivocally be assigned to either the *Z. v. carniolica* or the *Z. v. vivipara* clades, without signatures of hybridization (Fig. 2). It is worth noting that, with *K* = 2, two viviparous individuals from Austria (sampled far from the region were the two forms overlap geographically) seemed to show a mixed contribution, but with *K* > 2, it becomes clear that the apparent
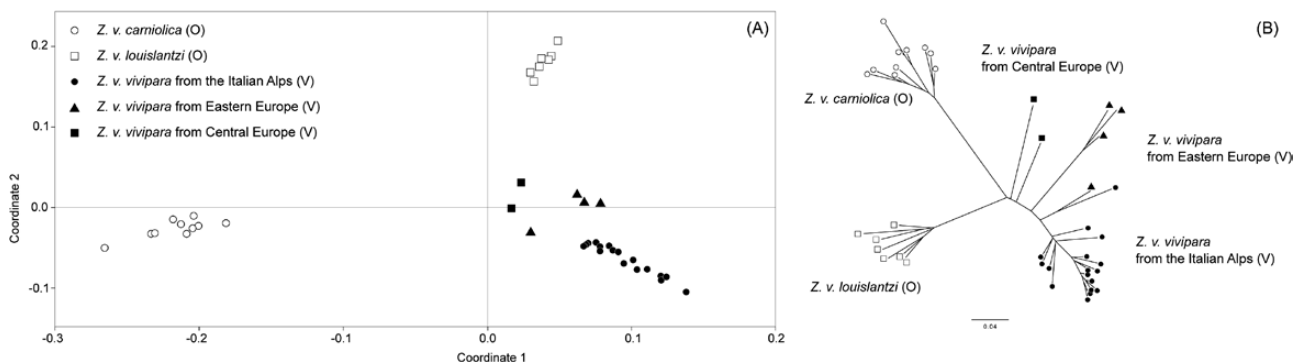

Figure 1. A, multidimensional scaling. B, unrooted neighbor-joining tree based on pairwise distances between individuals computed from a matrix of 80 792 SNPs. Abbrev: O, oviparous; V, viviparous.
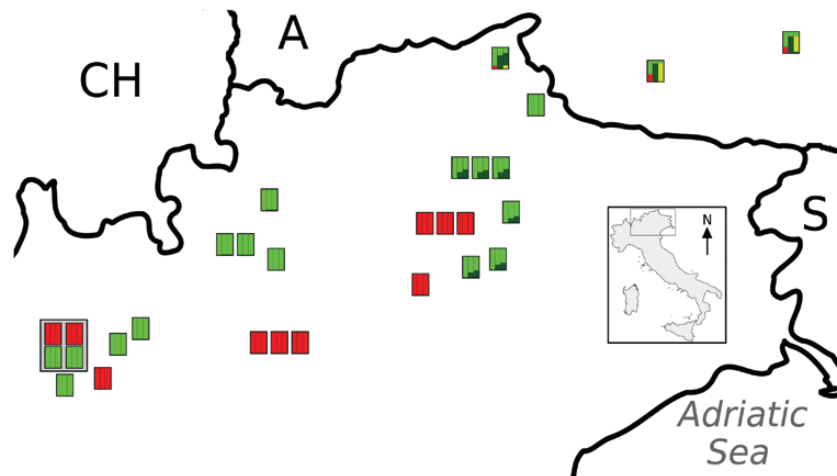
Figure 2. STRUCTURE results [based on 1403 single-nucleotide polymorphisms (SNPs) with < 30% of missing data] and geographic position of 29 Alpine individuals. Each individual is represented by a rectangle composed of three vertical bars corresponding to its genomic composition when *K* was fixed to 2, 3 and 4, respectively. Four individuals in the Western area are reported within a grey square to indicate that they come from a location where *Zootoca vivipara vivipara* and *Zootoca vivipara carniolica* live in sintopy and were sampled few metres apart from each other. The results are identical allowing 20% of missing data (602 usable SNPs).

oviparous ancestry found for *K* = 2 corresponds actually to a second viviparous component. Finally, when the Treemix approach was applied to all the five major groups, a tree with zero edges (i.e. no admixture) explained 99.85% of the variance. When a single migration edge was allowed, this was found to join two geographically and genetically close viviparous populations (Alps and Central Europe), and the variance explained by the model reaches 100%. In other words, a model without any introgression between viviparous and oviparous groups is able to entirely explain the genetic data.

GENES WITH SNPS CORRELATED WITH THE
SWITCH IN REPRODUCTIVE MODE

The analysis aimed at the identification of genes associated with reproductive mode was performed on a restricted set of 4908 SNPs with less than 50% of missing data. This analysis extracted 217 SNPs in 175 loci (Supporting Information, Fig. S2). A set of 34 loci were identified as putative homologs of an *A. carolinensis* gene using the DNA (BLAST) approach, and 33 loci were identified as putative homologs of an *A. carolinensis* gene using the protein (tBLASTx) approach. The c-list included 22 genes (genes identified by both approaches), and the i-list included 45 genes (genes identified by at least one approach). All these genes are listed in Supporting Information, Table S2. It is important to note that the two approaches use different data and should not be considered as two alternative methods to find the same set of SNPs. For example, relevant

genes without RAD loci in the coding region cannot be found by the protein-based approach, and large divergence at synonymous sites can prevent mapping when the DNA-based approach is used. In other words, even if genes in c-list are directly associated with SNPs in the coding region that safely map both at the DNA and at the protein level, this list probably fails to identify other relevant genes. We therefore believe that also the genes included in the i-list, but not in the c-list, should be evaluated with attention.

The biological role of the candidate genes is rather heterogeneous (see Supporting Information, Table S3), with a range of GO terms that include cytoskeleton organization, protease function, immune system processes, metabolic processes, transport, protein folding, cell adhesion, signalling and regulation of transcription. The small number of loci with a large variety of functions is probably responsible for the observed lack of significance in overrepresentation of GO terms considering both the c-list and i-list (see Supporting Information, Table S4 for the overrepresentation analysis performed on the c-list). Nevertheless, ~27 and 29% of the genes in the c-list and i-list, respectively, can be classified as transcription regulators, and 9% of the genes in the i-list are protease enzymes (Table 1).

DISCUSSION

This study is the first analysis of the genomic differences between populations of the common lizard *Z. vivipara* with different reproductive modes across

Table 1. List of the 12 genes with regulatory functions observed among the 45 genes (the i-list) that contain polymorphisms associated with the reproductive mode in *Zootoca vivipara*

ENSEMBL ID	Gene symbol	Gene name	Function description
ENSACAG00000022140	*TFE3**	TRANSCRIPTION FACTOR E3	This gene encodes a basic helix-loop-helix domain-containing transcription factor that binds MUE3-type E-box sequences in the promoter of genes.
ENSACAG00000010833	*EIF4E2**	EUKARYOTIC TRANSLATION INITIATION FACTOR 4E TYPE 2	It recognizes and binds the 7-methylguanosine-containing mRNA cap during an early step in the initiation of protein synthesis and facilitates ribosome binding by inducing the unwinding of the mRNAs secondary structures
ENSACAG00000008719	*SOX9**	TRANSCRIPTION FACTOR SOX9	The protein encoded by this gene recognizes the sequence CCTTGAG along with other members of the HMG-box class DNA-binding proteins
ENSACAG00000013802	*DACH2*†	DACHSHUND HOMOLOG 2	The encoded protein may be involved in regulation of organogenesis and myogenesis, and may play a role in premature ovarian failure
ENSACAG00000014366	*BACH2*†	TRANSCRIPTION REGULATOR PROTEIN BACH2	Transcriptional regulator that acts as repressor or activator
ENSACAG00000017962	*KAT2A**	HISTONE ACETYLTRANSFERASE KAT2A	Histone acetyltransferase that functions primarily as a transcriptional activator
ENSACAG00000002423	*SMARCA2**	GLOBAL TRANSCRIPTION ACTIVATOR SNF2L2-RELATED	Members of this family have helicase and ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin structure around those genes
ENSACAG00000011931	LOC100563157*	RNA BINDING PROTEIN FOX1 HOMOLOG 1	This gene encodes an RNA binding protein that is thought to be a key regulator of alternative exon splicing in the nervous system and other cell types
ENSACAG00000006538	*CLUH*†	CLUSTERED MITOCHONDRIA PROTEIN HOMOLOG	This gene binds mRNAs of nuclear-encoded mitochondrial proteins in the cytoplasm and regulates transport or translation of these transcripts close to mitochondria, playing a role in mitochondrial biogenesis
ENSACAG00000012332	*SNAI2*†	SNAIL FAMILY ZINC FINGER 2	Transcriptional repressor that modulates both activator-dependent and basal transcription
ENSACAG00000011132	*NOTCH1*‡	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1	Notch signalling is an intercellular signalling pathway that regulates interactions between adjacent cells through binding of Notch family receptors to their cognate ligands
ENSACAG00000017318	*WIZ*‡	PROTEIN WIZ (WIDE INTERSPERSED ZINC FINGER)	This gene is involved in protein and metal ion binding

The function description is reported according to GeneCards (www.genecards.org).
*The gene was identified with both the BLAST and the tBLASTx approach.
†The gene was identified only with the BLAST approach.
‡The gene was identified only with the tBLASTx approach.

its distributional range. Three main genetic clusters can be clearly distinguished, corresponding to the two oviparous subspecies *Z. v. carniolica* and *Z. v. louislantzi* and to the viviparous *Z. v. vivipara*. This pattern reflects the species genetic diversity, which has been shaped by geography and differential reproductive mode. In addition, gene flow between the subspecies *Z. v. vivipara* and *Z. v. carniolica*, which live sympatrically in the Alps, is very unlikely, at least in recent times, as previously suggested (Cornetti *et al.*, 2015b).

We conservatively identified 22 annotated genes putatively associated with different reproductive modes based on their pattern at the coding sequence and their significant mapping (both at protein and DNA level) to the reference genome of *A. carolinensis*. Additional 23 genes were found when mapping was based only on the DNA or on the protein approach. These genes are associated with diverse biological processes, which probably reflect the complex changes to physiology and morphology that occur in the transition from oviparity to viviparity. However, a considerable fraction of these candidate genes (almost 30%, independently of how candidate genes are identified) are involved in transcriptional regulation. We suggest that these genes play an important role in the transition between oviparity and viviparity, and we discuss this point in the following section.

REPRODUCTIVE MODE AND GENE EXPRESSION

We identified 12 candidate genes putatively involved in transcriptional regulation (Table 1). Changes to the protein sequence of regulators of transcription (e.g. transcription factors and cofactors) can result in changes to both the expression and diversity of target genes (Hsia & McGinnis, 2003). The physiological effects of changes to regulatory proteins will depend on where these proteins are expressed, how the polymorphism alters the amino acid sequence and the targets that the regulators interact with. Of our 12 candidate genes classified as transcriptional regulators, at least three genes are regulators of uterine development in eutherians. Their identification as candidates producing parity mode differences by our RAD-seq analysis further supports the involvement of these transcriptional regulators in the evolution of reproductive mode in *Z. vivipara*.

The evolution of viviparity is likely to involve physiological changes to three distinct tissues: the uterus, the ovary and extra-embryonic membranes. We used data from the Human Protein Atlas (Uhlén *et al.*, 2015) to identify whether our candidate regulators of transcription are expressed in human female reproductive tissues: either the endometrium, ovary, or trophoblast (referred to as the placenta in the Human Protein Atlas). In humans, the trophoblast forms from the chorioallantoic membrane, a tissue homologous

to that forming the embryonic component of the placenta in *Z. vivipara*. All 12 transcriptional regulators exhibited protein localization in one of these tissue types, except *EIF4E2*. Furthermore, all 12 regulators are also expressed in the uterus and chorioallantoic membrane of both oviparous and viviparous skinks (Griffith *et al.*, 2016a, b, 2017), suggesting that these regulators may be ancestrally expressed in the female amniote reproductive tract and extra-embryonic membranes. Given that the majority of these genes localize to reproductive tissues in mammals and are expressed in the squamate uterus and embryonic placental tissues, any modification to their protein sequence, which results in changes in protein function, will alter the gene expression landscape, potentially altering the reproductive physiology of *Z. vivipara*.

One of the striking genomic level changes observed in viviparous lizards is that they exhibit complex changes in uterine gene expression during pregnancy and appear to have a new pregnancy-specific state of gene regulation (Griffith & Wagner, 2017; Griffith *et al.*, 2016b). Therefore, it is exciting that our study has identified a significant number of transcription factors that are expressed in the uterus of both viviparous and oviparous lizards. Furthermore, three of these transcription factors are known to have important regulatory functions in the uterus of mammals during pregnancy, and we discuss them in detail: *DACH2*, *SOX9* and *NOTCH1*.

The Dachshund Family Transcription Factor 2 (*DACH2*) is a transcription co-factor with highly conserved protein interacting domains. Knockouts of *DACH2* in female mice result in developmental failure of the female reproductive tract (Davis *et al.*, 2008). Knockouts in males produce no similar defect, suggesting that the gene is important for the specialization of the female reproductive tract in mammals. In *Drosophila*, Dachshund, an ortholog of the human *DACH2* gene, facilitates development of the male and female reproductive tract (Keisman & Baker, 2001), suggesting that the function of this gene is broadly conserved across bilateral animals. Furthermore, *DACH2* is expressed in the oviduct of both viviparous (*P. entrecasteauxii*) and oviparous scincid lizards (*L. bougainvillii* and *Lampropholis guichenoti*; Griffith *et al.*, 2016b).

SOX9 is another putative transcriptional regulator identified in our study. In mammals, transcription factor *SOX9* is important for the development of the male phenotype. While *SOX9* is crucial for male development, the protein is also important for the development of glandular epithelium in the human endometrium (Gonzalez, 2012). *SOX9* is expressed in the uterine tissue of both gravid and non-reproductive oviparous and viviparous skinks, and may therefore also be important for squamate uterine development (Brandley *et al.*, 2012; Griffith *et al.*, 2016b).

NOTCH1 is a cell surface receptor in the *NOTCH* signalling pathway and is important for regulating cell-fate determination. *NOTCH1* is a key regulator of endometrial remodelling during pregnancy. Specifically, *NOTCH1* is essential for the maintenance of endometrial stromal cells and decidualization of the uterus (PrabhuDas *et al.*, 2015). Decreased *NOTCH1* production is associated with endometriosis in women (Su *et al.*, 2015). Uterine remodelling is probably a fundamental process for viviparous amniotes as it is necessary for the increased uterine vasculature necessary for embryonic gas exchange (Griffith & Wagner, 2017). The role of *NOTCH1* in *Z. vivipara* during reproduction requires further characterization but, as in mammals, evolution of the *NOTCH1* gene may allow for changes in gene expression associated with the remodelling of the uterus for pregnancy in viviparous lizards. Together, these results suggest that *DACH2*, *SOX9* and *NOTCH1* are likely to function as regulators of uterine development in *Z. vivipara*. Changes in the protein sequence of these genes may result in changes in uterine gene regulation, potentially providing the regulatory changes necessary to support the evolution of viviparity.

Our data set identifies SNPs in key regulatory genes that are correlated with changes in reproductive mode. These candidate genes could support the gene expression changes required for the evolution of viviparity, which have been shown in other studies (Brandley *et al.*, 2012; Griffith *et al.*, 2016b). However, more work is needed to test whether this correlation is causative; in particular, transcriptomic studies are needed to identify changes in gene regulation consistent with changes in reproductive mode, followed by functional studies to test the impacts of the identified SNPs in *Z. vivipara* tissues *in vitro*.

REPRODUCTIVE MODE AND ENZYMES

A final comment is dedicated to the candidate genes that may be involved in the evolution of viviparity by modifying the properties of enzymes in reproductive tissues. In fact, about 10% of the genes found in the most inclusive list (4 out of 45) encoded protease enzymes. Proteases are enzymes that cleave proteins at specific amino acid sites inside cells. These enzymes play an important role in the development of the uterus and placental tissues during pregnancy in both mammals and reptiles (Salamonsen & Nie, 2002; Song, Spencer & Bazer, 2005; Song *et al.*, 2010; Brandley *et al.*, 2012; Griffith *et al.*, 2016b) and are also critical for processes of pregnancy in viviparous fishes including seahorses (Whittington *et al.*, 2015b). In the viviparous skink *C. ocellatus*, the expression of Cathepsin L accounts for 5% of total uterine transcription during

pregnancy (Brandley *et al.*, 2012). The association of protease genes with lizard reproductive mode and the expression of proteases in the amniote uterus suggest that changes to the chemical properties of proteases may result in functional modification of these enzymes. We propose that the three protease genes expressed in the uterus or extra-embryonic membranes of *Z. vivipara* support changes to uterine morphology during pregnancy.

CONCLUSIONS

The evolution of viviparity from oviparity requires a suite of physiological changes to reproductive tissues that facilitate the internal incubation of embryos. Our results support the idea that these physiological changes are underpinned by modifications at a correspondingly extensive suite of genes, making the multiple origins of viviparity in vertebrates (Blackburn, 2015) even more remarkable. Our identification of regulator genes involved in the transition between reproductive modes confirms recent evidence of large-scale regulatory changes involved in the evolution of mammalian and reptile pregnancy (Griffith *et al.*, 2016b; Lynch *et al.*, 2016).

While our RAD-seq approach has identified genes associated with parity mode in *Z. vivipara*, these findings represent a correlation rather than a detailed identification of causal genetic changes involved in transitions between reproductive modes. This work provides the foundation for functional analyses of the candidate *Z. vivipara* genes identified here, which will ultimately be required to determine their mode of action and potential role in driving the evolution of viviparity.

REFERENCES

Adams SM, Biazik J, Stewart RL, Murphy CR, Thompson MB. 2007. Fundamentals of viviparity: comparison of seasonal changes in the uterine epithelium of oviparous and viviparous *Lerista bougainvillii* (Squamata: Scincidae). *Journal of Morphology* **268:** 624–635.

Angelini F, Ghiara G. 1984. Reproductive modes and strategies in vertebrate evolution. *Bolletino di Zoologia* **51:** 121–203.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3:** 1–7.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57:** 289–300.

Blackburn DG. 2015. Evolution of vertebrate viviparity and specializations for fetal nutrition: a quantitative and qualitative analysis. *Journal of Morphology* **276:** 961–990.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23:** 2633–2635.

Brandley MC, Young RL, Warren DL, Thompson MB, Wagner GP. 2012. Uterine gene expression in the live-bearing lizard, *Chalcides ocellatus*, reveals convergence of squamate reptile and mammalian pregnancy mechanisms. *Genome Biology and Evolution* **4:** 394–411.

Braz HB, Scartozzoni RR, Almeida-Santos SM. 2016. Reproductive modes of the South American water snakes: a study system for the evolution of viviparity in squamate reptiles. *Zoologischer Anzeiger: A Journal of Comparative Zoology* **263:** 33–44.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10:** 421.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22:** 3124–3140.

Cornetti L, Belluardo F, Ghielmi S, Giovine G, Ficetola GF, Bertorelle G, Vernesi C, Hauffe HC. 2015a. Reproductive isolation between oviparous and viviparous lineages of the Eurasian common lizard *Zootoca vivipara* in a contact zone. *Biological Journal of the Linnean Society* **114:** 566–573.

Cornetti L, Ficetola GF, Hoban S, Vernesi C. 2015b. Genetic and ecological data reveal species boundaries between viviparous and oviparous lizard lineages. *Heredity* **115:** 517–526.

Cornetti L, Menegon M, Giovine G, Heulin B, Vernesi C. 2014. Mitochondrial and nuclear DNA survey of *Zootoca vivipara* across the eastern Italian Alps: evolutionary relationships, historical demography and conservation implications. *PLoS ONE* **9:** e85912.

Davis RJ, Harding M, Moayedi Y, Mardon G. 2008. Mouse Dach1 and Dach2 are redundantly required for Müllerian duct development. *Genesis* **46:** 205–213.

Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

Gonzalez G. 2012. *Role of SOX9 in uterine gland development and disease initiation*. Unpublished Thesis, University of Texas, Houston, MD Anderson Cancer Center.

Griffith OW, Blackburn DG, Brandley MC, Van Dyke JU, Whittington CM, Thompson MB. 2015. Ancestral state reconstructions require biological evidence to test evolutionary hypotheses: a case study examining the evolution of reproductive mode in squamate reptiles. *Journal of Experimental Zoology B: Molecular and Developmental Evolution* **324:** 493–503.

Griffith OW, Brandley MC, Belov K, Thompson MB. 2016a. Allelic expression of mammalian imprinted genes in a matrotrophic lizard, *Pseudemoia entrecasteauxii*. *Development Genes and Evolution* **226:** 79–85.

Griffith OW, Brandley MC, Belov K, Thompson MB. 2016b. Reptile pregnancy is underpinned by complex changes in uterine gene expression: a comparative analysis of the uterine transcriptome in viviparous and oviparous lizards. *Genome Biology and Evolution* **8:** 3226–3239.

Griffith OW, Brandley MC, Whittington CM, Belov K, Thompson MB. 2017. Comparative genomics of hormonal signaling in the chorioallantoic membrane of oviparous and viviparous amniotes. *General and Comparative Endocrinology* **244:** 19–29.

Griffith OW, Wagner GP. 2017. The placenta as a model for understanding the origin and evolution of vertebrate organs. *Nature Ecology & Evolution* **1:** 72.

Guillaume CP, Heulin B, Pavlinov IY, Semenov D, Bea A, Vogrin N, Surget-Groba Y. 2006. Morphological variation in the common lizard, *Lacerta (Zootoca) vivipara*. *Russian Journal of Herpetology* **13:** 1–10.

Guillette L. 1982. The evolution of viviparity and placentation in the high elevation, Mexican lizard *Sceloporus aeneus*. *Herpetologica* **38:** 94–103.

Heulin B, Stewart JR, Surget-Groba Y, Bellaud P, Jouan F, Lancien G, Deunff J. 2005. Development of the uterine shell glands during the preovulatory and early gestation periods in oviparous and viviparous *Lacerta vivipara*. *Journal of Morphology* **266:** 80–93.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6:** e1000862.

Holsinger KE. 2000. Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 7037–7042.

Hsia CC, McGinnis W. 2003. Evolution of transcription factor function. *Current Opinion in Genetics & Development* **13:** 199–206.

Keisman EL, Baker BS. 2001. The *Drosophila* sex determination hierarchy modulates wingless and decapentaplegic signaling to deploy dachshund sex-specifically in the genital imaginal disc. *Development* **128:** 1643–1656.

Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, Emera D, Sheikh SZ, Grützner F, Bauersachs S, Graf A, Young SL, Lieb JD, DeMayo FJ, Feschotte C, Wagner GP. 2016. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. *Cell Reports* **10:** 551–561.

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. 2008. Database indexing

for production MegaBLAST searches. *Bioinformatics* **24:** 1757–1764.

Murphy BF, Thompson MB. 2011. A review of the evolution of viviparity in squamate reptiles: the past, present and future role of molecular biology and genomics. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* **181:** 575–594.

Paulesu L, Bigliardi E, Paccagnini E, Ietta F, Cateni C, Guillaume CP, Heulin B. 2005. Cytokines in the oviparity/viviparity transition: evidence of the interleukin-1 system in a species with reproductive bimodality, the lizard *Lacerta vivipara*. *Evolution & Development* **7:** 282–288.

Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics* **8:** e1002967.

PrabhuDas M, Bonney E, Caron K, Dey S, Erlebacher A, Fazleabas A, Fisher S, Golos T, Matzuk M, McCune JM, Mor G, Schulz L, Soares M, Spencer T, Strominger J, Way SS, Yoshinaga K. 2015. Immune mechanisms at the maternal-fetal interface: perspectives and challenges. *Nature Immunology* **16:** 328–334.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Qualls CP, Shine R. 2006. *Lerista bougainvillii*, a case study for the evolution of viviparity in reptiles. *Journal of Evolutionary Biology* **11:** 63–78.

Rawn SM, Cross JC. 2008. The evolution, regulation, and function of placenta-specific genes. *Annual Review of Cell and Developmental Biology* **24:** 159–181.

R Core Development Team. 2015. *R: a language and environment for statistical computing*. ISBN 3-900051-07-0. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org.

Rodríguez-Díaz T, Braña F. 2012. Altitudinal variation in egg retention and rates of embryonic development in oviparous *Zootoca vivipara* fits predictions from the cold-climate model on the evolution of viviparity. *Journal of Evolutionary Biology* **25:** 1877–1887.

Salamonsen LA, Nie G. 2002. Proteases at the endometrial-trophoblast interface: their role in implantation. *Reviews in Endocrine & Metabolic Disorders* **3:** 133–143.

Sandrock C, Schirrmeister BE, Vorburger C. 2011. Evolution of reproductive mode variation and host associations in a sexual-asexual complex of aphid parasitoids. *BMC Evolutionary Biology* **11:** 348.

Shine R. 2005. Life-history evolution in reptiles. *Annual Review of Ecology, Evolution, and Systematics* **36:** 23–46.

Sillero N, Campos J, Bonardi A, Corti C, Creemers R, Crochet PA, Isailovi JC, Denoël M, Ficetola GF, Gonçalves J, Kuzmin S, Lymberakis P, de Pous P, Rodriguez A, Sindaco R, Speybroeck J, Toxopeus B, Vieites DR, Vences M. 2014. Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-Reptilia* **35:** 1–31.

Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J, Arnaiz O, Awedh MH, Baldock R, Barbiera G, Bardou P, Beck T, Blake A, Bonierbale M, Brookes AJ, Bucci G, Buetti I, Burge S, Cabau C, Carlson JW, Chelala C, Chrysostomou C, Cittaro D, Collin O, Cordova R, Cutts RJ, Dassi E, Genova AD, Djari A, Esposito A, Estrella H, Eyras E, Fernandez-Banet J, Forbes S, Free RC, Fujisawa T, Gadaleta E, Garcia-Manteiga JM, Goodstein D, Gray K, Guerra-Assuncao JA, Haggarty B, Han DJ, Han BW, Harris T, Harshbarger J, Hastings RK, Hayes RD, Hoede C, Hu S, Hu ZL, Hutchins L, Kan Z, Kawaji H, Keliet A, Kerhornou A, Kim S, Kinsella R, Klopp C, Kong L, Lawson D, Lazarevic D, Lee JH, Letellier T, Li CY, Lio P, Liu CJ, Luo J, Maass A, Mariette J, Maurel T, Merella S, Mohamed AM, Moreews F, Nabihoudine I, Ndegwa N, Noirot C, Perez-Llamas C, Primig M, Quattrone A, Quesneville H, Rambaldi D, Reecy J, Riba M, Rosanoff S, Saddiq AA, Salas E, Sallou O, Shepherd R, Simon R, Sperling L, Spooner W, Staines DM, Steinbach D, Stone K, Stupka E, Teague JW, Dayem Ullah AZ, Wang J, Ware D, Wong-Erasmus M, Youens-Clark K, Zadissa A, Zhang SJ, Kasprzyk A. 2015. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research* **43:** 589–598.

Smith SA, Shine R. 1997. Intraspecific variation in reproductive mode within the Scincid lizard *Saiphos equalis*. *Australian Journal of Zoology* **45:** 435–445.

Song G, Bailey DW, Dunlap KA, Burghardt RC, Spencer TE, Bazer FW, Johnson GA. 2010. Cathepsin B, cathepsin L, and cystatin C in the porcine uterus and placenta: potential roles in endometrial/placental remodeling and in fluid-phase transport of proteins secreted by uterine epithelia across placental areolae. *Biology of Reproduction* **82:** 854–864.

Song G, Spencer TE, Bazer FW. 2005. Cathepsins in the ovine uterus: regulation by pregnancy, progesterone, and interferon tau. *Endocrinology* **146:** 4825–4833.

Stewart JR. 2013. Fetal nutrition in lecithotrophic squamate reptiles: toward a comprehensive model for evolution of viviparity and placentation. *Journal of Morphology* **274:** 824–843.

Stewart JR, Ecay TW, Heulin B. 2009. Calcium provision to oviparous and viviparous embryos of the reproductively bimodal lizard *Lacerta* (*Zootoca*) *vivipara*. *The Journal of Experimental Biology* **212:** 2520–2524.

Stewart JR, Ecay TW, Heulin B, Fregoso SP, Linville BJ. 2011. Developmental expression of calcium transport proteins in extraembryonic membranes of oviparous and viviparous *Zootoca vivipara* (Lacertilia, Lacertidae). *The Journal of Experimental Biology* **214:** 2999–3004.

Su RW, Strug MR, Joshi NR, Jeong JW, Miele L, Lessey BA, Young SL, Fazleabas AT. 2015. Decreased Notch pathway signaling in the endometrium of women with endometriosis impairs decidualization. *The Journal of Clinical Endocrinology and Metabolism* **100:** E433–E442.

Surget-Groba Y, Heulin B, Guillaume CP, Puky M, Semenov D, Orlova V, Kupriyanova L, Ghira I, Smajda B. 2006. Multiple origins of viviparity, or reversal from viviparity to oviparity? The European common lizard (*Zootoca vivipara*, *Lacertidae*) and the evolution of parity. *Biological Journal of the Linnean Society* **87:** 1–11.

Surget-Groba Y, Heulin B, Guillaume CP, Thorpe RS, Kupriyanova L, Vogrin N, Maslak R, Mazzotti S, Venczel M, Ghira I, Odierna G, Leontyeva O, Monney JC, Smith N. 2001. Intraspecific phylogeography of *Lacerta vivipara* and the evolution of viviparity. *Molecular Phylogenetics and Evolution* **18:** 449–459.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research* **13:** 2129–2141.

Thompson MB, Speake BK. 2006. A review of the evolution of viviparity in lizards: structure, function and physiology of the placenta. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* **176:** 179–189.

Touchon JC, Warkentin KM. 2008. Reproductive mode plasticity: aquatic and terrestrial oviposition in a treefrog. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 7495–7499.

True JR, Carroll SB. 2002. Gene co-option in physiological and morphological evolution. *Annual Review of Cell and Developmental Biology* **18:** 53–80.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CAK, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F. 2015. Tissue-based map of the human proteome. *Science* **347:** 126419.

Van Dyke JU, Brandley MC, Thompson MB. 2014. The evolution of viviparity: molecular and genomic data from squamate reptiles advance understanding of live birth in amniotes. *Reproduction* **147:** R15–R26.

Whittington CM, Grau GE, Murphy CR, Thompson MB. 2015a. Unusual angiogenic factor plays a role in lizard pregnancy but is not unique to viviparity. *Journal of Experimental Zoology B: Molecular and Developmental Evolution* **324:** 152–158.

Whittington CM, Griffith OW, Qi W, Thompson MB, Wilson AB. 2015b. Seahorse brood pouch transcriptome reveals common genes associated with vertebrate pregnancy. *Molecular Biology and Evolution* **32:** 3114–3131.

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics* **42:** 355–360.

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44:** 821–824.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Details of samples analysed in this study. Mitochondrial clade according to Surget-Groba *et al.* (2006). In addition, number of retained reads and mean coverage per sample are reported.

Table S2. Alignment statistics of the 45 genes (the i-list) identified as homologs in *Anolis carolinensis* genome and possibly involved in the switch in reproductive mode. Number of mismatches, length of the alignment, length of the contig and the e-value of the alignment are reported for both single-end and the mini-contigs generated using paired-end reads. The total alignment length is also reported. Genes are grouped according to their function. Genes putatively involved in transcriptional regulation have a yellow background, proteases have a grey background and genes with other functions have a white background. Genes producing both a BLAST and a tBLASTx mapping with the *Anolis carolinensis* genome (i.e., genes belonging to the c-list) are in bold.

Table S3. List of genes putatively associated to the switch in reproductive modality in *Zootoca vivipara* with their biological function, as reported by PantherDB using the *Anolis carolinensis* reference genome. Genes are grouped according to their function. Genes putatively involved in transcriptional regulation have a yellow background, proteases have a grey background and genes with other functions have a white background. Genes producing both a BLAST and a tBLASTx mapping with the *Anolis carolinensis* genome (i.e., genes belonging to the c-list) are in bold.

Table S4. Resume of the overrepresentation test performed with Panther on the i-list. Significant enrichments (P value < 0.05, without Bonferroni correction) are listed with increasing P values. For each GO biological process the number of genes in *Anolis carolinensis* and the number of genes among the candidate from the conservative list observed in *Zootoca vivipara* are reported, as well as the expected number of genes associated with each category and the estimated fold enrichment. None of the P values is significant after multiple test correction (as implemented in Panther), also when the c-list is analysed.

Figure S1. Distributions of Weir & Cocherham estimates of F_{st} in different pairwise comparisons. SNPs within the red bins in all the three comparisons were identified as candidate markers responsible for the switch in reproductive mode under the 'F_{st} based' method (see main text for details). Zvv = *Z. v. vivipara*; Zvc = *Z. v. carniolica*; Zvl = *Z. v. louislantzi;* V = Viviparous; O = Oviparous.

Figure S2. Venn diagram reporting the overlap between the SNPs under selection identified by three different methods.