An endogenous retroviral envelope syncytin and its cognate receptor identified in the viviparous placental *Mabuya* lizard

Guillaume Cornelis[a,b,1,2], Mathis Funk[a,b,1], Cécile Vernochet[a,b], Francisca Leal[c,3], Oscar Alejandro Tarazona[c,4], Guillaume Meurice[d], Odile Heidmann[a,b], Anne Dupressoir[a,b], Aurélien Miralles[e], Martha Patricia Ramirez-Pinilla[c], and Thierry Heidmann[a,b,5]

[a]Unité Physiologie et Pathologie Moléculaires des Rétrovirus Endogènes et Infectieux, CNRS UMR 9196, Gustave Roussy, Villejuif, F-94805, France; [b]UMR 9196, Université Paris-Sud, Orsay, F-91405, France; [c]Laboratorio de Biologia Reproductiva de Vertebrados, Escuela de Biologia, Universidad Industrial de Santander, 680002 Bucaramanga, Colombia; [d]Plateforme de Bioinformatique, INSERM US23/CNRS UMS3655, Gustave Roussy, Villejuif, F-94805, France; and [e]Institut de Systématique, Evolution, Biodiversité, Muséum National d'Histoire Naturelle, CNRS UPMC EPHE, Sorbonne Universités, Paris, F-75005, France

Edited by R. Michael Roberts, University of Missouri-Columbia, Columbia, MO, and approved October 26, 2017 (received for review August 23, 2017)

Syncytins are envelope genes from endogenous retroviruses that have been captured during evolution for a function in placentation. They have been found in all placental mammals in which they have been searched, including marsupials. Placental structures are not restricted to mammals but also emerged in some other vertebrates, most frequently in lizards, such as the viviparous *Mabuya* Scincidae. Here, we performed high-throughput RNA sequencing of a *Mabuya* placenta transcriptome and screened for the presence of retroviral *env* genes with a full-length ORF. We identified one such gene, which we named "*syncytin-Mab1*," that has all the characteristics expected for a syncytin gene. It encodes a membrane-bound envelope protein with fusogenic activity ex vivo, is expressed at the placental level as revealed by in situ hybridization and immunohistochemistry, and is conserved in all *Mabuya* species tested, spanning over 25 My of evolution. Its cognate receptor, required for its fusogenic activity, was searched for by a screening assay using the GeneBridge4 human/Chinese hamster radiation hybrid panel and found to be the MPZL1 gene, previously identified in mammals as a signal-transducing transmembrane protein involved in cell migration. Together, these results show that syncytin capture is not restricted to placental mammals, but can also take place in the rare nonmammalian vertebrates in which a viviparous placentotrophic mode of reproduction emerged. It suggests that similar molecular tools have been used for the convergent evolution of placentation in independently evolved and highly distant vertebrates.

endogenous retrovirus | envelope protein | placenta | syncytin | receptor

Syncytins are "captured" genes of retroviral origin that correspond to the envelope gene of ancestrally endogenized retroviruses. These genes encode fusogenic proteins that are involved in the formation, by cell–cell fusion, of the syncytiotrophoblast at the placental materno–fetal interface in eutherian mammals (reviewed in refs. 1 and 2). Furthermore, genetically modified mice in which the two syncytin genes, *syncytin-A* and *syncytin-B*, were knocked out showed deficiencies in placental development, with altered structure of the materno–fetal interface resulting in inhibition of growth or death of the embryo at midgestation (3, 4). Syncytins have been found in all placental mammals in which they have been searched, with independently captured syncytins occurring across all major clades of placental mammals, including Euarchontoglires (primates, rodents, and lagomorphs), Laurasiatherians (ruminants and carnivores), Afrotherians (tenrec), and even Marsupials (opossum) (Fig. 1) (5–15). This has led to the proposal that these genes, which are absolutely required for placentation as shown by the knockout mice experiments, are most likely involved in the emergence and evolution of placental mammals from egg-laying animals (1). Analysis of the conservation of these genes further indicates that they have been subjected to purifying selection in the course of evolution, as expected for any bona fide cellular gene.

Remarkably, placental structures are not restricted to mammalian species. Placentation emerged independently and in a stochastic manner in several groups of vertebrates, with the noticeable exception of birds (reviewed in refs. 16 and 17). In particular, complex placentas have been described in some South American or African species of Scincidae lizards (18–21; reviewed in refs. 16, 22, and 23). In one Scincidae genus, *Mabuya*, placental structures form specialized regions very similar to those found in mammals, e.g., the placentome where maternal and invasive fetal tissues are highly folded and interdigitated. Ultrastructural analysis of this region further revealed the presence of a syncytial structure at the materno–fetal interface, as observed in numerous mammalian species. However, in contrast to most mammals, the syncytial layer of the *Mabuya* placenta is formed by maternal epithelial cells

Significance

Retroviral envelope gene capture and exaptation for a placental function has been demonstrated in mammals. Remarkably, placental structures have also emerged on rare occasions in nonmammalian vertebrates, resulting in related modes of reproduction. The *Mabuya* lizard, which emerged 25 Mya, possesses a placenta closely related to that of mammals. Here, we identified a specific retroviral envelope gene capture that shows all the characteristic features of a bona fide mammalian syncytin, being conserved in *Mabuya* evolution, expressed in the placenta, and fusogenic. Together with the present identification of its cognate receptor, these results show that syncytin capture is not restricted to mammals and is likely to be a major driving force for placenta emergence.

Author contributions: G.C., M.F., and T.H. designed research; G.C., M.F., and C.V. performed research; F.L., O.A.T., A.M., and M.P.R.-P. collected and processed live biological materials and samples; G.M. contributed analytic tools; G.C., M.F., C.V., G.M., O.H., A.D., A.M., M.P.R.-P., and T.H. analyzed data; and G.C., M.F., and T.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: RNA sequencing data are deposited in the European Nucleotide Archive, www.ebi.ac.uk/ena (accession no. ERA1116158). Mabuya sequences described in this paper (Mabuya *MPZL1*, *syncytin-Mab1*, and *Mab-Env2–Mab-Env4*) have been deposited in the GenBank database (accession nos. MG254887–MG254891, respectively).

See Commentary on page 13315.

[1]G.C. and M.F. contributed equally to this work.

[2]Present address: Department of Genetics, Stanford University, Stanford, CA 94305.

[3]Present address: Department of Biology, University of Florida, Gainesville, FL 32611.

[4]Present address: Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL 32611.

[5]To whom correspondence should be addressed. Email: heidmann@igr.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1714590114/-/DCSupplemental.

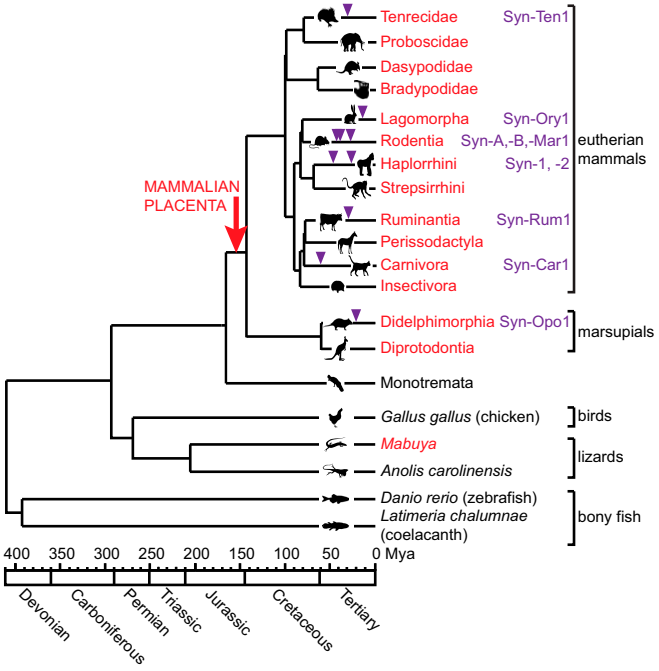

Fig. 1. Phylogeny of vertebrates positioning mammals, the *Mabuya* lizard, and known syncytins. Mammals comprise the monotremes (e.g., platypus) still laying eggs, and the marsupials and eutherian mammals, which all possess a placenta (red font). The lizard *Mabuya* is shown in red font because it also possesses a placenta. A red arrow indicates the probable time of emergence of the mammalian placenta, which has been proposed to correspond to the primitive capture of an ancestral syncytin, thereafter replaced in evolution by the indicated present-day syncytins (reviewed in ref. 1). All currently described syncytin capture events are indicated by arrowheads in purple, together with the syncytin's name. Branch length is proportional to time [expressed in My (15, 64, 65)], as indicated in the scale below the tree.

rather than by fetal trophoblast cells, which remain individualized cells throughout gestation (19, 20, 24–27). Yet these differences can be considered minor in view of the tremendous qualitative transition that must have taken place between egg-laying and placental lizards that have a mode of reproduction so closely related to that of placental mammals. Following our previous proposal that such transitions are likely due to the stochastic acquisition of new genes via the endogenization of retroviruses and the exaptation of their envelope gene, a process that has taken place on several occasions in mammalian evolution, we investigated whether, as in mammals, endogenous retroviral *env* genes could have been specifically captured and exapted for a role in placenta formation in the *Mabuya* genus and might even be involved in the refined structure of their materno–fetal interface. Such events would indeed represent a remarkable example of convergent evolution in two major classes of vertebrates.

Currently, no genome within the *Mabuya* genus has been sequenced, and the most closely related species whose genome is available is the green anole (*Anolis carolinensis*), a distant small oviparous lizard. As a result, we could not screen *Mabuya* genomes to identify candidate syncytin genes as was done previously for mammalian species. After having captured pregnant females from a Colombian *Mabuya* population [referred to as "*Mabuya* sp. IV" (28)], we therefore performed high-throughput RNA sequencing (RNA-seq) of their placenta transcriptome and searched for the presence of expressed, coding retroviral *env* genes. Remarkably, we identified such a gene, *syncytin-Mab1*, which proved to be a syncytin and was characterized for its fusogenic activity, its conservation during *Mabuya* evolution, and its expression at the materno–fetal interface. Using a genetic approach, we identified its cognate receptor as the MPZL1 transmembrane protein, previously characterized as a membrane-bound signal transducer implicated in cell-migration processes (29–32). This identification of a captured syncytin and its receptor in a distant viviparous lizard with a placentotrophic mode of reproduction shows that syncytin capture is not restricted to mammals and is likely to be a major driving force for placenta emergence and evolution.

Results

High-Throughput Sequencing and in Silico Search for Retroviral *Env* Genes Within the *Mabuya* Placental Transcriptome. Placental RNA was extracted from pregnant Andean *Mabuya* sp. IV females captured in the wild (*Methods*). The transcriptome of a placenta at embryonic development stage 35 (as defined in ref. 33), presenting a well-established syncytium, was determined by the French National Sequencing Center (Genoscope, Evry, France). Placental transcripts were assembled using de novo assembly methods in absence of a reference genome (*Methods*). Transcriptome reconstruction identified 72,763 transcripts ranging from 123 bp to 24.4 kb [N50 (the length N for which 50% of all bases in the sequences are in a sequence of length $L < N$) = 2,869 bp; median size = 1.1 kb], 26,690 of which (36.7%) could be positively matched to a known protein gene from the Refseq Vertebrate database. A set of 21,253 transcripts shows significant expression (more than two tags per million) (34), 16,398 of which were positively annotated (77.2%) and correspond to 13,189 unique genes, among which 10,079 were nonhousekeeping genes. This subset of transcripts contains a range of genes annotated as "placental" genes with key biological functions during placenta development, such as core placental transcription factors, placental hormones, iron transporters, gap junction proteins, interleukins and interleukin receptors, metalloproteinases, and angiogenesis factors (*SI Transcriptome Analysis* and Datasets S2–S4).

To identify putative *env*-derived syncytin genes in this transcriptome, we screened these assemblies using the method we previously devised to screen mammalian genomes for such genes. In short, a BLAST search for *env* ORFs longer than 400 aa (from the Met start codon to the stop codon) was performed using a set of *env* sequences including all previously identified syncytins (*Methods*). It yielded four *Mabuya* transcripts that we named "*Mab-Env1*" to "*Mab-Env4*." Analysis of the overall structure of the four identified *env* genes (Fig. 2 *A* and *B*) strongly suggests that they indeed correspond to retroviral Env proteins, with all or most of their characteristic features conserved, including the presence of a predicted signal peptide sequence at the N terminus, a putative furin/PACE cleavage site delineating a surface (SU) and a transmembrane (TM) subunit, and a hydrophobic domain >20 aa long located downstream of the highly conserved C-X_{5-7}-C motif of retroviral envelopes. This hydrophobic domain is involved in anchoring the Env protein within the plasma membrane, a feature required for its fusogenic activity. Finally, all genes present a canonical immunosuppressive domain (ISD). The presence of the ISD classifies the four envelopes as gamma-type retroviral envelopes (reviewed in ref. 35); however only Mab-Env3 and Mab-Env4 present the gamma-type–specific C-X_6-CC motif that contains the third cysteine necessary for the covalent link between SU and TM observed in gamma-type envelopes, with the other two cysteines forming an intra-TM loop. Mab-Env1 and Mab-Env2 present a C-X_7-C motif, a characteristic of beta-type retroviral envelopes in which the SU and TM are not linked by a disulfide bond and only the intra-TM bond is present (reviewed in ref. 35).

Incorporation of Mab-Env1 to Mab-Env4 into a phylogenetic tree based on the alignment of TM subunits is shown in Fig. 2*C* and demonstrates that they are distinct from previously identified *env* genes and that all four of them cluster with gamma-type envelopes. Interestingly, Mab-Env1 and Mab-Env2 are grouped together and display 82% amino acid identity with strong conservation of most of the functional domains (Fig. 2*D*), suggesting that they correspond to two genomic copies of the same retroviral family. The main differences between Mab-Env1 and Mab-Env2 are a 10-aa deletion in the N-terminal domain of the TM subunit of Mab-Env2 (the domain presumably harboring the fusion peptide critical for Env fusion activity) and a short truncation of the

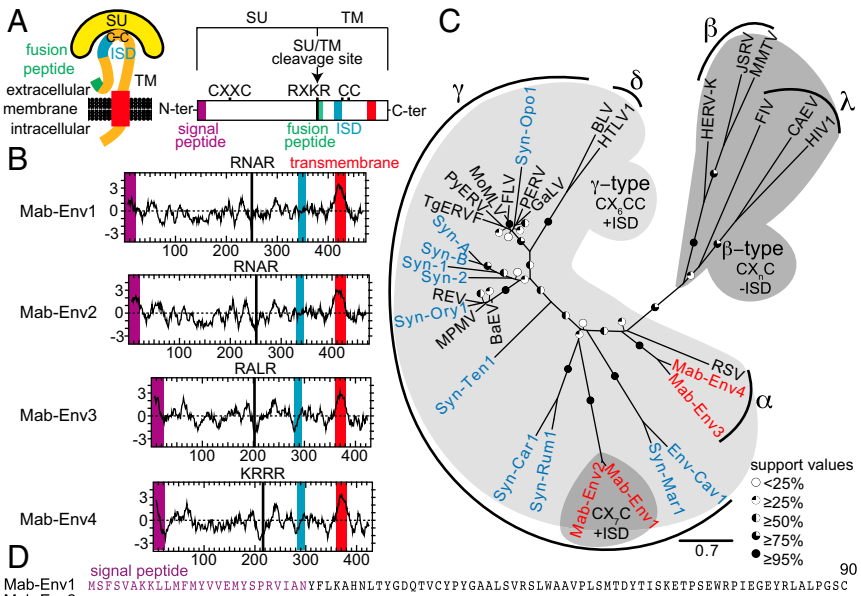

Fig. 2. Structure of a canonical retroviral Env protein and characterization of the identified *Mabuya* candidates. (*A*) Schematic representation of a retroviral Env protein delineating the SU and TM subunits. The furin cleavage site (consensus: R/K-X-R/K-R) between the two subunits, the C-X-X-C motif involved in SU–TM interaction, the hydrophobic signal peptide (purple), the fusion peptide (green), the transmembrane domain (red), the putative immunosuppressive domain (ISD) (blue), and the conserved C-X_{5-7}-C motif (CC) are indicated. (*B*) Characterization of the candidate *Mabuya* Env proteins. The hydrophobicity profile for each candidate is shown with the canonical structural features highlighted in *A*. (*C*) Retroviral envelope protein-based phylogenetic tree with the identified Mab-Env protein candidates in red. The unrooted PhyML maximum likelihood tree was obtained using TM subunit amino acid sequences (without the cytoplasmic tail) from syncytins (in blue) and a series of endogenous and infectious retroviruses (from the dataset in ref. 66). The horizontal scale bar below the tree represents the average number of substitutions per site, and the percent bootstrap values obtained from 1,000 replicates are indicated by circles on each branch (see key in figure). Retroviral families and major envelope types, the associated C-X_n-(C)C motif, and the presence or absence of an ISD are indicated on the tree. (*D*) Amino acid sequences and characteristic structural features of *Mab-Env1*/*Mab-2* and *Mab-Env3*/*Mab-4* (GenBank accession nos. MG254888–MG254891). Asterisks indicate amino acid identity; colons indicate amino acid similarity.

cytoplasmic tail of Mab-Env1 due to the presence of a premature stop codon. Although Mab-Env3 and Mab-Env4 are also grouped together, they only display 56% amino acid identity, suggesting that they belong to distinct, but related, retroviral families. Surprisingly, Mab-Env3 and Mab-Env4 cluster with the envelope gene of the Rous sarcoma virus (RSV), an alpharetrovirus presenting an avian gamma-type envelope. Closer inspection further reveals that their fusion peptide is located about nine amino acids after the putative furin site, a characteristic of avian gamma-type envelopes (reviewed in ref. 35). As their name indicates, this subset of gamma-type envelopes was thought to be specific to bird retroviruses.

***Mab-Env1* Is Highly Expressed in Placental Tissues.** We performed an RT-qPCR analysis of transcript levels for each candidate *Mab-Env* gene in different tissues (Fig. 3) using primers that were designed to be specific for each *env* sequence (Table S1). *Mabuya* placental RNA was analyzed at two developmental stages: stage 30, when the uterine cells start to fuse, forming the syncytial layer, and stage 37, when the syncytium is well formed. High expression levels of *Mab-Env1* (similar to that of the ribosomal protein RPL19) could be observed in the placenta at both stages. The expression of *Mab-Env2*, *Mab-Env3*, and *Mab-Env4* in the placenta was 6- to 85-fold lower than that of *Mab-Env1*. Of note, none of these genes was specifically expressed in the placenta. *Mab-Env1*, in particular, shows high levels of expression in several other organs, including organs of the maternal genital tract (ovary and oviduct), a feature not commonly observed for mammalian syncytins. Due to the high expression levels of *Mab-Env1* in the placenta and the low expression levels of the three other *env* genes, we focused on *Mab-Env1* in the present study.

Characterization of the *Mab-Env1*–Associated *Pol* Gene. To obtain the sequence of the provirus harboring the *Mab-Env1* gene, we rescreened the placental transcriptome in an iterative fashion using BLAST and found additional overlapping transcripts, each sharing a 3′ region homologous with the 5′ region of the previous transcript, allowing us to reconstruct an internal proviral sequence by concatenating the transcripts (Fig. 4*A*). We confirmed the existence of this sequence through PCR on genomic DNA with primers designed at both extremities of the assembled RNA-seq sequence (Table S1). We were able to amplify and sequence 5.8 kb of internal proviral sequence containing the end of *gag* and the

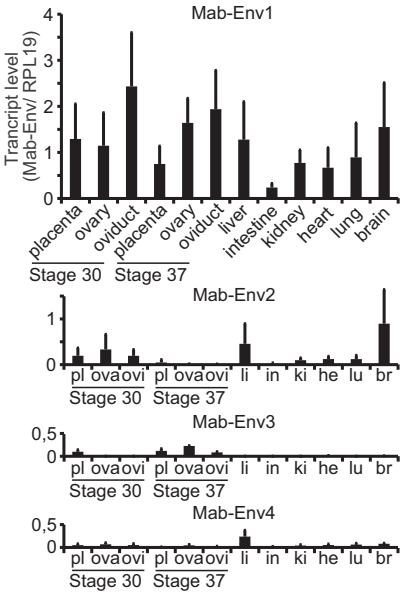

Fig. 3. Real-time RT-qPCR analysis of candidate *env* gene expression in *Mabuya* sp. IV. Transcript levels are shown as the ratio between the expression levels of each *env* gene and those of the *RPL19* control gene (*SI Methods*). Placental tissues and maternal ovary and oviduct tissues were recovered at two gestational stages (defined using ref. 33): stage 30, when the uterine syncytium starts to form in the placentome region, and stage 37, when the syncytium is well developed. The results for the four *env* candidate genes were obtained using the same series of tissues. Values are the means of duplicates from three samples ± SEM.

entirety of *pol* and *Mab-Env1*. The *pol* ORF is interrupted at the 3′ end by several stop codons and frameshifts, while the *Mab-Env1* ORF is intact, as expected (Fig. 4*A*). Of note, an acceptor splice site can be predicted with a high score in the proviral sequence, about 400 bp upstream of the *env* start codon (using www.cbs.dtu.dk/services/NetGene2), but the corresponding retroviral spliced *env* transcript could not be retrieved from the RNA-seq assemblies.

Identification of the conserved reverse transcriptase (RT) domain of the *pol* gene allowed us to compare it with previously described RT domains to determine the family of the *Mab-Env1*–associated virus. A PhyML maximum likelihood tree using RTs from both exo- and endogenous retroviruses (Fig. 4*B*) places the *Mab-Env1*–associated RT closest to the gamma family of RTs. However, it remains external to all gammaretroviral RTs of the dataset, indicating that it might be a distant member of this family or a member of another, closely related family. In association with the noncanonical combination of an ISD and a C-X_7-C motif in *Mab-Env1* (see above and Fig. 2*C*), this finding suggests that this lizard retrovirus differs from previously described retroviral families.

Conservation Within the *Mabuya* Genus. To study *Mab-Env1* conservation within the *Mabuya* genus, we tentatively PCR-amplified the entire *Mab-Env1* ORF from the genomic DNA of several species of Scincidae. Fragments of the expected length were obtained for all *Mabuya* species shown in Fig. 5. Sequencing of the purified PCR products revealed intact reading frames belonging to the *Mab-Env1* family (Fig. 5). No amplification was obtained in any non-*Mabuya* Scincidae, even though amplification of *BDNF*, a conserved control gene, indicated that the extracted genomic DNA was of good quality. We then used a less stringent approach by designing degenerate primers in conserved internal *Mab-Env1* regions based on alignments of all available *Mabuya* sequences. This approach led to the amplification of a 400-bp fragment, ranging from the end of the SU to the start of the ISD, in all *Mabuya* tested, as expected, but also in the *Trachylepis*, *Chioninia*, *Eumecia*, and *Lubuya* genera (sequences are shown in Fig. S2). This unexpected result indicates that endogenization of the *Mab-Env1*–associated virus probably occurred before the separation of these genera, about 30 Mya (Fig. 5) (36), but that the full-length envelope was conserved only in the *Mabuya* genus. Of note, we have been unable to amplify even the 400-bp fragment from *Heremites vitattus* genomic DNA, which is not compatible with the entry date proposed above (Fig. 5). *Heremites* is a newly created genus (37), and its position on the Scincidae phylogenetic tree is only weakly supported so far, which might explain this apparent anomaly, although sequence divergence could also account for this negative result.

In Situ Analyses of *Mab-Env1* Transcription and Protein Expression on Placental Sections. As described previously in refs. 19, 20, and 24–26, the *Mabuya* placenta is a complex structure composed of specialized regions (Fig. 6*A*) including areolae and absorptive plaques involved in histiotrophic nutrition and, at the embryonic pole, the placentome and paraplacentome. In the placentome region, maternal and fetal epithelia fold and interdigitate during gestation, and ramifications are highly vascularized. At the materno–fetal interface, the fetal side is mainly composed of giant columnar binucleated chorionic cells. On the maternal side, the uterine epithelium is replaced by large syncytial structures, most probably resulting from the cellular fusion of uterine cells. At the cellular interface, maternal and fetal cells form microvilli, further increasing the exchange surface of the placentome (19, 20, 24–26). Periodically the microvilli are interrupted by the presence of

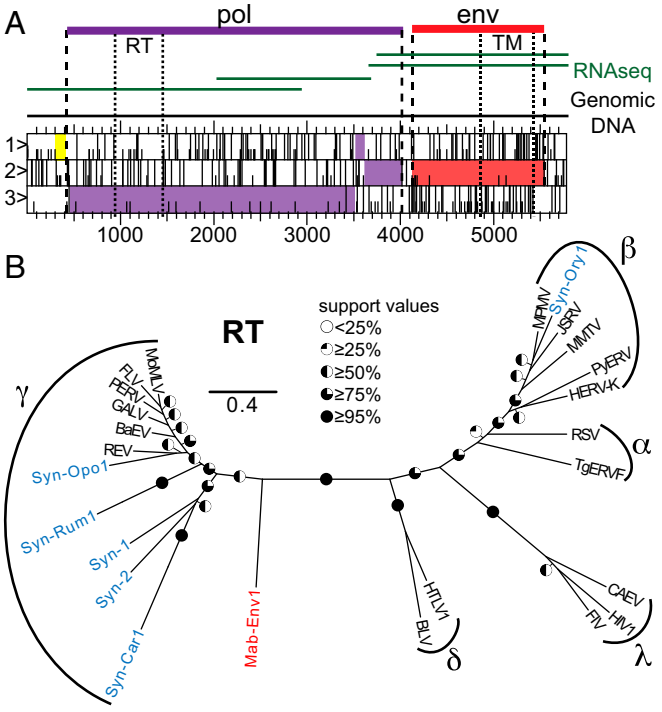

Fig. 4. Characterization of the *Mab-Env1*–associated RT-containing *pol* gene and classification among known retroviral families. (*A*) ORF map of the identified genomic proviral fragment. Short bars indicate a start codon; tall bars indicate a stop codon. Green horizontal lines represent overlapping RNA-seq transcripts. Similarities to canonical retroviral genes by BLASTp search are colored in yellow for *gag* and purple for *pol*. *Mab-Env1* is colored red. The conserved RT and TM domains used for the phylogenetic analyses in Figs. 2*C* and 4*B* are delineated. (*B*) Unrooted PhyML tree showing the position of retroviral endogenous and exogenous RT domains using the same dataset as in Fig. 2*C* when available (some syncytin-associated RTs were too degenerate to be identified). The *Mab-Env1*–associated RT is highlighted in red, and those associated with previously described syncytins are shown in blue. Major retroviral families are indicated on the tree. Bootstraps obtained after 1,000 replicates are indicated by circles on each branch (see key in figure). Branch length is proportional to the average number of substitutions per site (see scale). RT sequences are provided as Dataset S1.

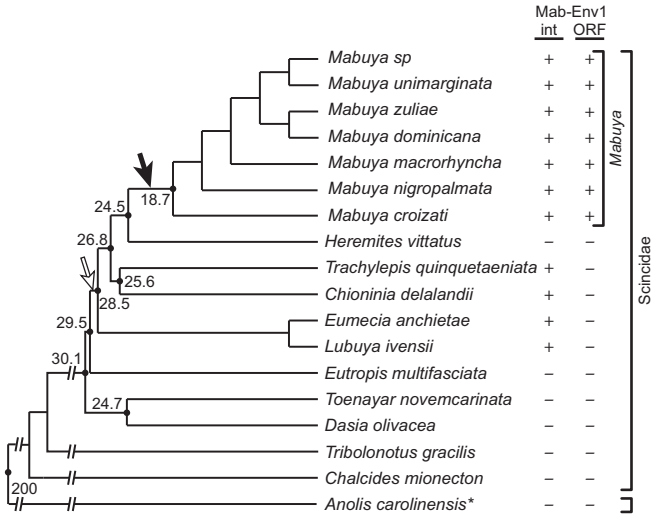

Fig. 5. *Mab-Env1* is conserved only within the *Mabuya* genus. (*Left*) Species tree of Scincidae with *A. carolinensis* as an outgroup. Nodes with black circles are dated in Mya and placed proportionally to their age (28, 36, 64). Entry of *Mab-Env1* is indicated by the white-tipped arrow; clades in which the envelope has been conserved are indicated by the black arrow. An asterisk indicates that a published genome is available (*A. carolinensis*). (*Right*) A plus sign indicates amplification of a *Mab-Env1* internal 400-bp fragment (int) or the complete *env* ORF by genomic DNA PCR. Amplification of the complete ORF is restricted to the *Mabuya* genus, with no amplification in other Scincidae.

smaller invasive fetal cells located between the fetal and maternal epithelia, allowing a more intimate contact between fetal tissues and the maternal circulatory system (38). The paraplacentome surrounds the placentome and is limited on one side by the abrupt transition from a uterine syncytium to a uterine epithelium. The organization of fetal tissues in the paraplacentome does not differ significantly from that observed in the placentome, although tissues are not folded. The materno–fetal interface is still marked by interdigitated microvilli, although they are smaller than in the placentome (19, 20, 24–26).

To assess the potential physiological relevance of *Mab-Env1* expression in the placenta, both in situ hybridization and immunohistochemistry experiments were performed on paraffin sections of stage 37 *Mabuya* placentas (Fig. 6*B*). Specific digoxigenin-labeled antisense riboprobes were synthesized for the detection of *Mab-Env1* transcripts, and the corresponding sense riboprobes were used as negative controls. Anti–Mab-Env1 sera were obtained by immunizing mice with Mab-Env1 SU, and preimmune sera were used as controls (*SI Methods*). Both methods show that the expression of *Mab-Env1* seems to be widespread, but the staining (at both the RNA and protein level) is more intense in the cell layers forming the interface between mother and fetus. In the placentome, the staining indicates stronger expression in the maternal syncytium (ms) than in the fetal chorionic cells (fc). In the paraplacentome, strong staining is observed at the apical side of the fetal chorionic cell layer, including the microvillous interface between mother and fetus. These expression profiles would support a physiological role for *Mab-Env1* at the materno–fetal interface (as further examined below).

***Mab-Env1* Is Fusogenic in ex Vivo Assays.** The functionality of Mab-Env1 as an ancestrally derived retroviral envelope protein was assayed as previously for other syncytins. First, we tested whether this envelope protein added *in trans* could render a recombinant retrovirus deprived of its native *env* gene able to infect target cells (Fig. 7*A*). To do so, the *env* gene was cloned and introduced into a CMV promoter-containing expression vector that was used to generate pseudotyped murine leukemia virus (MLV) particles (*SI Methods*). As illustrated in Fig. 7 *B* and *C*, Mab-Env1–pseudotyped MLV particles generated in human 293T cells could infect a panel of target cells of different origins, including human (293T, TE671, SHSY-5Y, HeLa), bovine (MDBK), carnivoran (G355.5, CrFK, MDCK), and rodent (WOP, 208F, A23) cells, in some cases at levels similar to those observed using the amphotropic MLV (A-MLV) Env protein (although all rodent cells tested exhibited a significantly lower titer). Since Mab-Env1 is a protein of reptilian origin, this suggests that its cognate receptor is a highly conserved protein among vertebrates. Rodent cells might be less permissive to infection due to subtle divergences in the evolution of the receptor in this rapidly evolving clade. Mab-Env2 was found to be negative for all target cells tested under the same experimental conditions. Although we cannot exclude the possibility that Mab-Env2 has lost its ability to interact with the mammalian ortholog of its cognate receptor or that Mab-Env1 and Mab-Env2 have different receptors, these are unlikely. Indeed, as described above, Mab-Env1 and Mab-Env2 are very similar, but Mab-Env2 presents a deletion within the putative fusion peptide (Fig. 2*D*), suggesting that Mab-Env2 has lost its fusogenic properties.

Cell–cell fusion experiments were performed by transfecting cells with the *env* expressing plasmids and assaying the formation of syncytia 48 h post transfection in the presence or absence of an acidic shock (Fig. 7*D*). Mab-Env1 was able to induce fusion only when the cells were submitted to a low-pH shock, and under these conditions it induced the formation of large, multinuclear syncytia (Fig. 7*E*) in all cell lines tested except rodent cells, as expected from the assay described in Fig. 7*B*. Mab-Env2 tested negative under all conditions, further confirming the results of the pseudotyping assay.

Of note, both Mab-Env3 and Mab-Env4 were negative in the two assays above.

Search for the Mab-Env1 Receptor. Since Mab-Env1 is able to efficiently mediate the infection of human cells but not hamster A23 cells, we took advantage of the human–hamster hybrid irradiation panel (Genbridge4; ref. 39) in which fragments of the human genome had been introduced into a panel of 93 hamster A23 clones (Fig. S3*A*) and which we had used successfully to identify the receptor of human *Syncytin-2* (40). We screened the Genebridge4 clones for infection using the pseudotyping assay described above. The infection results for the 93 ordered clones were [0000010000001110000000001000000010000001010000000001000000001011000000200000000001], where 1 denotes infection (>100 focus-forming units/mL), 0 denotes the absence of infection, and 2 indicates that the clone was unavailable for analysis (Fig. S3*B*). This pattern of infection was compared with a matrix of specific genomic markers by computing a matching score using the RH program (41). The eight best-scoring markers were located within the same 2-Mb region of the human genome, on chromosome 1 (Fig. S3*C*). Among the genes located within this genomic region, 11 encode putative transmembrane proteins. The capacity of each of these genes to mediate the infection of A23 cells by Mab-Env1–pseudotyped particles, and thus to function as a receptor for Mab-Env1, was assayed using expression vectors for the corresponding cloned human cDNAs (Thermo Fisher) and revealed that only the myelin protein zero-like 1 (MPZL1) protein, a previously characterized single-pass transmembrane receptor (29, 42), was able to increase infection of A23 cells (see below). To verify that Mab-Env1 can also interact with *Mabuya* MPZL1, we searched by BLAST for the corresponding cDNA sequence in the *Mabuya* placental transcriptome and designed primers to amplify MPZL1 from placental cDNA. The *Mabuya* MPZL1 ORF that could be cloned shares 65% amino acid identity with human MPZL1 (Fig. 8*A*), with the glycosylation sites, the disulfide bond predicted for human MPZL1, and the intracellular C-terminal domain that contains two immunoreceptor tyrosine-based inhibitory motifs (ITIMs) (Fig. 8 *A* and *B*) (42) being conserved, whereas the extracellular N-terminal domains of the two proteins are much more variable. As illustrated in Fig. 8*C*, when infected with Mab-Env1–pseudotyped MLV particles, A23 cells transfected with either human or *Mabuya* MPZL1 exhibit titers 100 times

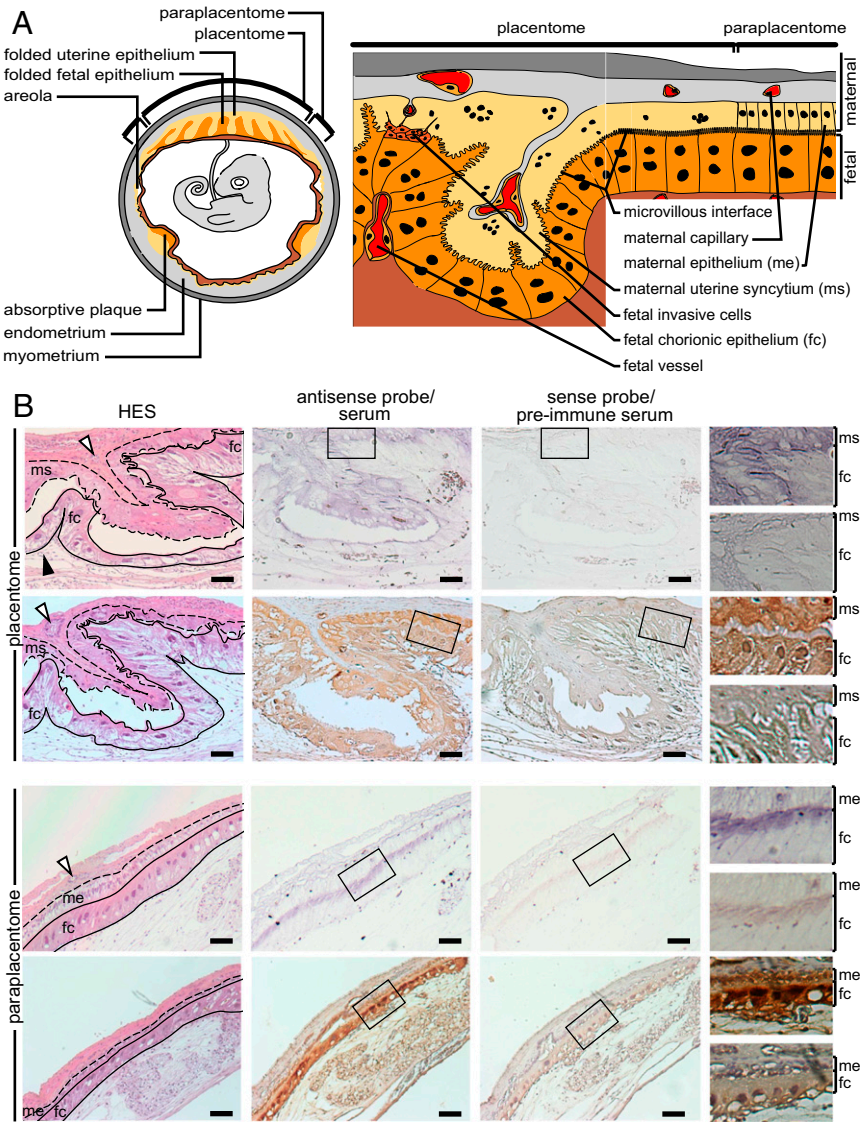

Fig. 6. Structure of the *Mabuya* sp. IV placenta and expression pattern of *Mab-Env1*. (*A*) Schematic representation of the late-stage *Mabuya* sp. IV placenta (see ref. 27). (*Left*) Overview of a gravid uterus displaying the apposed maternal and fetal tissues. Maternal and fetal tissues are highly interdigitated in the placentome region. In addition to the placentome, the *Mabuya* placenta develops specialized regions for materno–fetal exchanges such as areolae and absorptive plaques. The brown/orange and gray/yellow areas represent the fetal and maternal tissues, respectively. (*Right*) Detailed scheme of the materno–fetal interface in the placentome and paraplacentome region. In the placentome, maternal and fetal tissues are highly interdigitated, with numerous microvilli between fetal and uterine cells. The fetal epithelium is formed by giant binucleated chorionic cells. The uterine epithelium is replaced by a large syncytial structure formed by the fusion of the uterine cells. In the paraplacentome the syncytium is abruptly replaced by uterine epithelium. Cells still present microvilli, but the tissues are no longer interdigitated. (*B, Left*) Hematoxylin eosin saffron (HES)-stained sections of placenta with the maternal uterine syncytium (ms), the maternal epithelium (me), and the fetal chorionic epithelium (fc) delineated; white arrowheads indicate maternal and black arrowheads indicate fetal blood vessels. (*B, Right*) In situ hybridization (odd rows) or immunohistochemistry (even rows) on serial sections for *Mab-Env1* expression using digoxigenin-labeled antisense and sense riboprobes or an anti–Mab-Env1 mouse serum and preimmune serum. (Scale bars: 50 μm.) Areas marked by rectangles are enlarged on the right, and the maternal and fetal domains are delineated.

higher than control cells, indicating that *Mabuya* MPZL1 can indeed act as a receptor for Mab-Env1.

Expression and Potential Activation of MPZL1 in *Mabuya* Placenta. To determine if MPZL1 can interact with Mab-Env1 in the placenta, we first compared its expression profile with that of *Mab-Env1*. Using the same cDNA panel as in Fig. 3, we performed RT-qPCR and found that MPZL1 is indeed expressed in the placenta but also in other tissues (Fig. 8*D*). We then performed in situ hybridization on *Mabuya* placenta sections, as in Fig. 6*B*, using MPZL1-specific antisense probes and sense probes as negative controls. As illustrated in Fig. 8*E*, a specific antisense-only signal is observed at the materno–fetal interface. The maternal syncytium in the placentome and the maternal epithelium of the paraplacentome in particular show strong staining, while fetal chorionic cells are only moderately stained in both regions. This pattern is close to that observed for Mab-Env1 (Fig. 6*B*), especially for the placentome where both genes are more expressed at the level of the maternal syncytium, whereas in the paraplacentome the strongest expression of each gene takes place in distinct, although interdigitated, tissues of the materno–fetal interface. Therefore Mab-Env1 and MPZL1 could interact physically in vivo in both the placentome and paraplacentome, but in the former this interaction could contribute directly to the establishment of the maternal syncytium, which shows strong staining for both genes.

Previous studies have shown that MPZL1 is a signal transducer, the activation of which requires phosphorylation of the two tyrosine residues (Y241 and Y263) contained in its cytoplasmic ITIM motifs (Fig. 8 *A* and *B*) (29). To further demonstrate the interaction of Mab-Env1 with MPZL1, we investigated whether Mab-Env1 can trigger MPZL1 phosphorylation. To do so, we transduced untransformed human breast MCF10A cells with a lentiviral vector expressing *Mab-Env1* or a negative control (empty vector or a vector expressing the mammalian *syncytin-Car1*). As a positive control, we treated the cells with concanavalin A (Con A), which is known to induce homodimerization and phosphorylation of MPZL1 (29). Cell lysates were then analyzed by Western blot using an antibody directed specifically against the phosphorylated form as well as an antibody detecting both forms of MPZL1. In three independent transduction experiments we observed that Mab-Env1 very significantly reduces the total amount of MPZL1 without decreasing (rather slightly increasing) its overall phosphorylation level, compared with the control *syncytin-Car1* or empty vector (Fig. 9). This indicates that Mab-Env1 interacts with MPZL1, triggering both its degradation (reminiscent of receptor interference) and, as expected, a strong correlative increase of its phosphorylation.

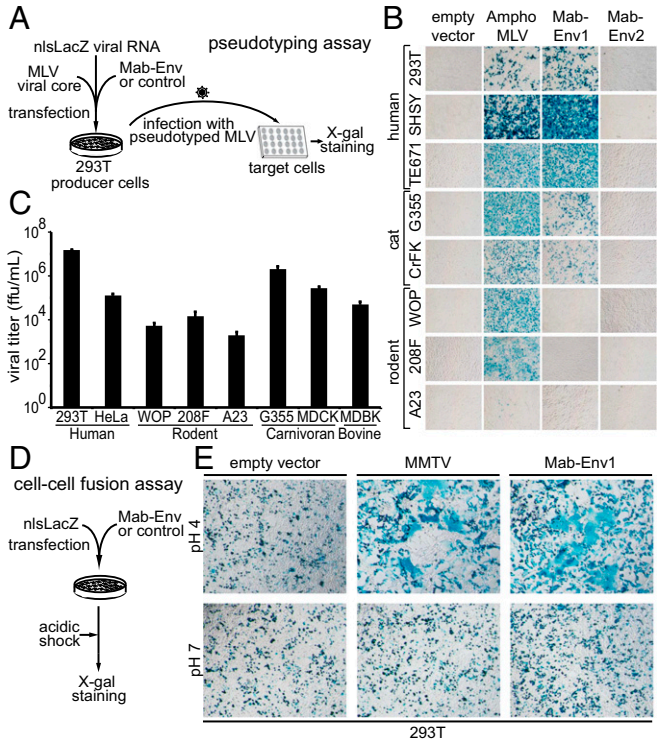

Fig. 7. Mab-Env1 is a fusogenic retroviral envelope protein. (*A*) Schematic representation of the cell infection assay with Mab-Env–pseudotyped virus particles. Pseudotypes are produced by cotransfecting 293T cells with expression vectors for the MLV core, a β-galactosidase encoded by a nls*LacZ*-containing retroviral transcript, and either a vector expressing Mab-Env proteins or a control vector. The supernatant of the transfected cells is then added to the indicated target cells, which are X-gal stained 3 d postinfection to reveal viral infection. (*B*) Panel of X-gal–stained target cells infected with particles without Env or pseudotyped with Mab-Env1, Mab-Env2, or the Env protein from A-MLV as positive control. (*C*) Quantification of viral titers expressed in focus-forming units per milliliter after infection with Mab-Env1–pseudotyped MLV virions. Values shown are the mean of three independent experiments ± SD. (*D*) Schematic representation of the cell–cell fusion assay. 293T cells were cotransfected with a plasmid expressing a nuclear β-galactosidase and a plasmid expressing either Mab-Env1 or a control plasmid. After 48 h a 5-min acidic shock was performed (or not) using Dulbecco's phosphate-buffered saline (DPBS)·HCl at either pH 4 or 7. Cells were X-gal stained 6 h afterward. (*E*) Mab-Env1 is able to induce cell–cell fusion after an acidic shock. After exposition to acid medium, cells transfected with the Mab-Env1–expressing plasmid show a degree of fusion equivalent to that of cells transfected with a plasmid expressing the Env protein of mouse mammary tumor virus (MMTV), which was used as a positive control.

Discussion

Analysis of the presently established *Mabuya* lizard placental transcriptome revealed the presence of an endogenous retroviral *env* gene, *Mab-Env1*. It is highly expressed in placental tissues, especially at the level of the materno–fetal interface, both in the maternal syncytium layer in the placentome and in the fetal epithelium in the paraplacentome. Mab-Env1 is a fusogenic protein that is able to functionally replace a present-day retroviral *env* gene within a recombinant infectious retrovirus. It also can mediate cell–cell fusion, although only under conditions of low pH as observed for the ruminant *syncytin-Rum1* (11). The cognate receptor of Mab-Env1 was also identified and found to be the previously described MPZL1 transmembrane signal transducer (29). Its expression is not restricted to the placenta, but in situ hybridization demonstrates expression at the materno–fetal interface, which colocalizes with that of *Mab-Env1* in the placentome syncytial layer. Finally, we have shown that *Mab-Env1* is conserved in the *Mabuya* species tested, covering ∼25 My of evolution of this specific placental clade (36). Together, these canonical characteristics—i.e., fusion activity, expression in the placenta, and conservation during evolution—allow *Mab-Env1* to be now named "*syncytin-Mab1*" and to be added to the series of previously characterized syncytins that we and others have identified so far in both Eutherian and Marsupial placental mammals (5–15).

The captured *Mab-Env1* and associated *pol* genes have some unusual properties that could be linked to the remoteness of the species in which they were discovered. Although the Env gene falls within the family of gamma-type envelopes according to both the phylogenetic tree and the presence of an ISD, as shown in Fig. 2, it possesses a C-X_7-C motif in the TM subunit that is not classically found in such envelopes. Gamma-type envelopes normally possess three cysteines in this motif (C-X_6-CC), allowing the formation of a covalent SU–TM link, whereas a two-cysteine motif is observed in betaretroviruses and lentiviruses, in which the SU and TM are not covalently associated (reviewed in ref. 35). The associated noncoding *pol* gene displays a characteristic RT domain that clusters more closely to gammaretroviral RTs but forms a sister clade to all remaining gammaretroviral sequences in our dataset (Fig. 4). This suggests that it is either a remotely related gammaretrovirus or that it belongs to another, potentially unknown, family. The latter might also explain the unusual characteristics of the *syncytin-Mab1* gene discussed above. It is also noteworthy that *Mab-Env3* and *Mab-Env4* present characteristics of avian gamma-type envelopes, further reinforcing the idea that the remoteness of the investigated species should reveal novel categories of endogenous retroviruses.

Other interesting results of the present investigation concern the nature of the identified receptor for this retroviral envelope and the properties of this previously characterized gene that could be relevant to *Mabuya* placenta physiology. The MPZL1 receptor is a single-pass transmembrane glycoprotein belonging to the immunoglobulin superfamily, with an intracellular moiety containing two ITIMs. Previous studies have shown that MPZL1 can promote the migration of mesenchymal-derived mouse embryonic fibroblasts (30, 31) or hepatocellular carcinoma cells (32) and can be involved in adhesion-dependent signaling (43, 44). In this respect, it will be interesting to determine whether the presently demonstrated *syncytin-Mab1*–induced phosphorylation of MPZL1 could also trigger receptor-mediated signaling in vivo that could participate in some function essential for *Mabuya* placentation in addition to the canonical membrane fusion activity observed for syncytins. For instance, activation of MPZL1 could enhance placental cell migration/invasion into the maternal tissue or even inhibit some specific immune functions via its ITIM domains.

One remarkable aspect of vertebrate physiology concerns the highly diverse modes of reproduction, with clades such as birds having a strictly oviparous mode of reproduction and other clades containing both oviparous and some viviparous species. Even among viviparous species, a range of reproductive strategies exists. In some species, the egg and embryo develop with very limited maternal supply, being essentially dependent on yolk supply (lecithotrophy), while at the other extreme some species show strong dependence on maternal exchanges and almost no reliance on yolk (matrotrophy) (reviewed in refs. 16 and 45). The placental organization of mammals represents the latter extreme, with the mother supplying almost all nutrients necessary for fetal growth. Clearly, the mammalian placenta is likely to be a monophyletic trait that emerged about 150 Mya and has since evolved into several variations on a common theme. In all cases in which it has been searched, there is evidence that retroviral capture has played a role in the diversity of placental structures, and possibly even in the emergence of the mammalian placenta via the capture of a primitive retroviral envelope during the radiation of this family (reviewed in refs. 1 and 2). Of note, there is increasing evidence that some nonmammalian vertebrates, among them *Mabuya* Scincidae and other lizards such as the African *Lubuya ivensii* (21, 46), possess a mode of viviparity with placental structures that closely resembles that found in mammals.

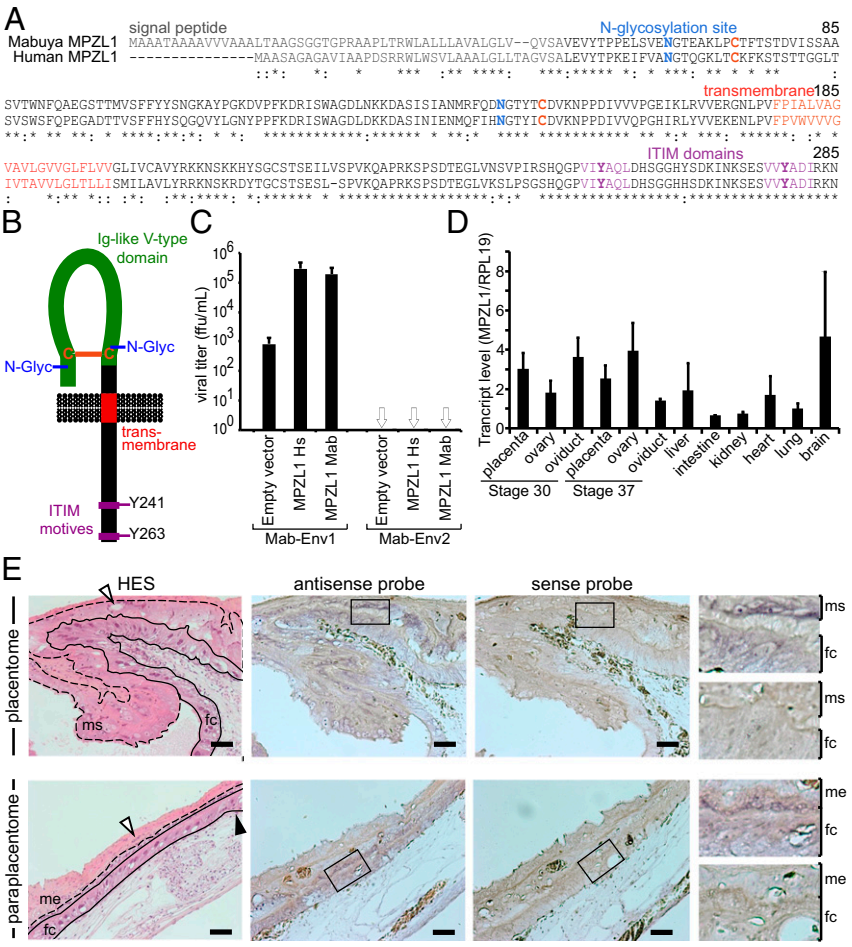

Fig. 8. Structure and expression of the Mab-Env1 receptor MPZL1. (*A*) Alignment of *Mabuya* MPZL1 (GenBank accession no. MG254887) and human MPZL1 amino acid sequences with major domains and sites indicated. Asterisks indicate amino acid identities, and colons indicate amino acid similarities. (*B*) To-scale schematic representation of the mature human MPZL1 with major domains indicated. (*C*) Human and *Mabuya* MPZL1 act as receptors for Mab-Env1–mediated infection. Target A23 cells were transfected with a plasmid expressing human or *Mabuya* MPZL1 or an empty vector. Twenty-four hours posttransfection, cells were infected using MLV particles carrying an nls*LacZ* reporter gene and pseudotyped with Mab-Env1 or Mab-Env2 or without Env. Cells were X-gal stained 72 h after infection, and infection foci were quantified. Values shown are the mean of three independent experiments ± SD. (*D*) MPZL1 is expressed in a wide range of tissues. RT-qPCR analysis of MPZL1 expression levels was performed in the same way and on the same series of tissues as in Fig. 3. (*E*) MPZL1 is expressed at the materno–fetal interface. (*Left*) HES-stained sections of placenta with the maternal syncytium (ms), the maternal epithelium (me), and the fetal chorionic epithelium (fc) delineated; white arrowheads indicate maternal and black arrowheads indicate fetal blood vessels. (*Right*) In situ hybridization on serial sections for *MPZL1* expression using digoxigenin-labeled antisense and sense riboprobes. (Scale bars: 50 μm.) Areas marked by rectangles are enlarged on the right, and the maternal and fetal domains are delineated.

These species have developed placentas that possess several of the characteristic features of the mammalian placenta and enable very similar reproductive strategies, including high maternal dependence for embryo growth, very tiny eggs almost devoid of yolk supply, prolonged pregnancy, and a large increase in conceptus dry mass during gestation (18, 22, 47). Accordingly, and taking the present results into account, we propose as a working hypothesis an extended version of a model previously proposed for mammals, in which emergence of this reproductive strategy is associated with the stochastic acquisition of a retroviral envelope gene that provides the corresponding species with syncytin functions.

This model can be substantiated as follows. First, it is now established that placenta-like structures have emerged among egg-laying vertebrates throughout vertebrate evolution, ranging from 400 Mya in fish to 150 Mya in mammals and 25 Mya in the *Mabuya* lizards (48; reviewed in refs. 17 and 49). These emergences appear to be random and rare. Consistently, retroviral endogenization events have been identified in a large series of vertebrates, with some of them having occurred over 400 Mya in marine species (50). Envelope gene capture in which this gene has been conserved after retrovirus endogenization has recently been characterized in spiny-rayed fishes dating back 100 My (51). This indicates that retroviral *env* capture has most probably taken place across all of vertebrate evolution, and endogenization probably even predates the emergence of vertebrates. Therefore, the stochastic emergence of specific modes of reproduction during vertebrate evolution could, in principle, be accounted for by retroviral capture events, taking into account the ancestry of these elements. One other hint provided by the present investigation is that we not only could demonstrate retroviral *env* capture in the form of the functional fusogenic envelope protein syncytin-Mab1 but also could correlate its capture in time with the emergence in the *Mabuya* genus of a placental structure ∼25 Mya (36). Although we acknowledge that the temporal correlation between the acquisition of a retroviral *env* gene possessing all the characteristic features of canonical mammalian syncytins and the emergence of the *Mabuya* placenta is not proof of a causal link, it constitutes a significant hint for a direct involvement (see also the properties of the *syncytin-Mab1*

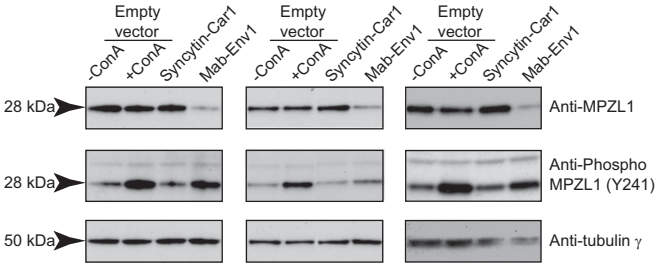

Fig. 9. *Mab-Env1* induces degradation and phosphorylation of MPZL1. MCF10A cells were transduced either with an empty lentiviral vector or with the same vector coding for *syncytin-Car1* or *Mab-Env1*. Con A treatment was performed (or not) 1 h before cell lysis as a positive control. Lysates were deglycosylated using PNGase F, and Western blot analysis was performed using an antibody recognizing both phosphorylated and unphosphorylated MPZL1 (*Top*) or only phosphorylated MPZL1 (*Middle*). (*Bottom*) An anti–tubulin-γ antibody was used to compare total protein levels in each well. Each set of blots represents an independent set of transduced cells.

receptor likely to be relevant to such a role, as discussed above). To be complete, we must also mention that part of the *syncytin-Mab1* sequence, although not the complete *env*, could be identified in two clades closely related to the *Mabuya* genus, namely the clade containing the *Trachylepis* and *Chioninia* genera and the clade containing the *Eumecia* and *Lubuya* genera, but not in any other, more distantly related clades (notably *Dasia*, *Toenayar*, or *Eutropis*). This suggests that *Mab-Env1* probably entered the Scincidae family slightly before the divergence of the *Mabuya* genus dated ∼25 Mya, and that the related *env* gene was not maintained as a protein-coding sequence in *Trachylepis* and *Chioninia* or in *Eumecia* and *Lubuya*. For the former two, this scenario would be consistent with *Trachylepis* being oviparous and *Chioninia* being lecithotrophic (52). For the latter two, which display matrotrophic placentation (21, 47, 52), it must be hypothesized either that the full-length gene could not be amplified for technical reasons or that another capture has taken place later in evolution. Along this line, it is theorized that the emergence of viviparity in American (*Mabuya*) and African (*Eumecia* and *Lubuya*) Scincidae results from two independent emergence events (52).

In any event, the identification in the present paper of both a syncytin and its cognate receptor in a remote nonmammalian species opens the door to further searches for other syncytin captures in distant placental vertebrates and to in-depth investigations of their roles. These could include noncanonical functions, as suggested here, that would add to their fusogenic activity. This could lead to a deeper understanding of the mechanisms underlying the rare and stochastic emergence of placentation, a remarkable example of evolutionary convergence, and provide hints on species evolution in vertebrates.

Methods

Animals and Tissues. *Mabuya* females were captured in the wild by M.P.R.-P. in the municipality of Curiti, Department of Santander, Colombia. This Colombian population is an as-yet unnamed species described in ref. 28 as *Mabuya* sp. IV. This species corresponds to the one studied in refs. 19, 24, 25, 27, and 53. Placental and uterine tissues were collected from killed pregnant females and were either stored in RNAlater (Qiagen) or processed for in situ hybridization. Other tissues (ovary, oviduct, liver, intestine, kidney, heart, lung, and brain) were collected from the same killed females. Total RNA was extracted after tissue lysis in TRIzol reagent (Invitrogen), and genomic DNA was extracted by phenol-chloroform extraction. Genomic DNA from *Mabuya croizati*, *Mabuya dominicana*, *Mabuya nigropalmata*, *Mabuya macrorhyncha*, *Mabuya unimarginata* (*sensu lato*), and *Mabuya zuliae* and from *Chalcides mionecton*, *Chioninia delalandii*, *Eutropis multifasciata*, *Heremites vittatus*, *Trachylepis quinquetaeniata*, and *Tribolonotus gracilis* was obtained by phenol-chloroform extraction from tissue samples provided by A.M. Similarly, genomic DNA was extracted from *Dasia olivacea* tissues provided by Aaron Bauer (Villanova University, Villanova, PA) and from *Lubuya ivensii* tissues (Museum of Comparative Zoology numbers R-193864, R-193868, R-193869, and R-193870) provided by Breda Zimkus (The Louis Agassiz Museum of Comparative Zoology, Cryogenic Collection, Harvard University, Cambridge, MA). Genomic DNA from *Eumecia anchietae* and *Toenayar novemcarinata* was also provided by Aaron Bauer.

***Mabuya* sp. IV Placenta Transcriptome High-Throughput Sequencing and de Novo Assembly.** Total RNA was purified from *Mabuya* sp. IV placenta before sequencing using the Illumina HiSeq. 2000 sequencing system. cDNA library construction, sequencing, and transcripts assembly were performed by the French National Sequencing Center (Genoscope, Evry, France). Briefly, poly(A)$^+$ RNA was selected with oligo(dT) beads, chemically fragmented, and converted into single-stranded cDNA using random hexamer priming according to the Illumina TruSeq protocol. The second strand was then generated to create double-stranded cDNA. Next, the paired-end library was prepared following Illumina's protocol: Briefly, fragments were end-repaired and then 3′-adenylated, and Illumina adapters were added by using NEBNext Sample Reagent Set (New England Biolabs); ligation products were purified, and DNA fragments (>200 bp) were PCR-amplified using Illumina adapter-specific primers. After library profile analysis (showing a typical library size of 200–600 bp) by Agilent 2100 Bioanalyzer (Agilent Technologies) and qPCR quantification (MxPro; Agilent Technologies), the library was sequenced using 100-base-length read chemistry in a paired-end flow cell on the Illumina HiSeq2000 instrument (Illumina). Nucleotide sequence information has been deposited at the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under the accession number ERA1116158. To maximize the number of reconstructed transcripts, several independent assemblies were performed using two different de novo assembly software programs (Oases and Trinity) as well as a range of k-mer lengths (55, 59, 63, 67, 71, 75 and 79), and only transcripts of >100 bp were conserved. We then used the EvidentialGene pipeline to remove duplicates and concatenate fragmented transcripts from each individual assembly (54). Reads were mapped on the nonredundant assembly using Bowtie2 (55), and transcript abundance was quantified using Express (56). To identify the transcripts, we used TBLASTX (word_size 3, gapopen 10, gapextend 3, BLOSUM 45) with an e-value of 10^{-3} to compare them with the nonredundant protein database of vertebrate genomes (RefSeq database Vertebrates January 5, 2015), and used the best hit as the transcript identifier for downstream analysis. Gene identifiers were converted to gene symbols using the bioDBnet (https://biodbnet-abcc.ncifcrf.gov) website for comparison with transcriptomes of other species [human and mouse placenta (57), *Chalcides ocellatus* pregnant uterus (58), and Eutherian uterine decidua (59)]. For functional annotations and enrichment, housekeeping genes were first removed (60), and functional annotations [GO Term, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and Mouse Genome Informatics Mammalian Phenotype] were performed using the Enrichr website (amp.pharm.mssm.edu/Enrichr/).

Database Screening and Sequence Analyses. Retroviral endogenous *env* gene sequences were searched in each transcriptome assembly. Sequences containing an ORF longer than 400 amino acids (from start to stop codons) were extracted from the transcriptomes using the getorf program of the EMBOSS package (emboss.sourceforge.net/apps/cvs/emboss/apps/getorf.html) and were translated into amino acid sequences. These sequences were compared with the TM subunit amino acid sequences of 35 retroviral envelope glycoproteins from representative endogenous retroviruses, among which are known syncytins, and infectious retroviruses using the BLASTp program (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Putative envelope protein sequences were then selected based on the presence of a hydrophobic domain (transmembrane domain) located 3′ to a highly conserved C-X_{5-7}-C motif. Multiple alignments of amino acid sequences were carried out using the MUSCLE algorithm in AliView (61). Maximum likelihood phylogenetic amino acid trees were constructed with PhyML 3.0 (62), with bootstrap percentages computed after 1,000 replicates. Substitution models were selected using the SMS program (63): LG+Γ+I for TM sequences and RtREV+ Γ for RT sequences.

Ethics Statement. This study was carried out in strict accordance with the French and European laws and regulations regarding animal experimentation (Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes). *Mabuya* were collected under the permit issued by the Ministerio de Ambiente y Desarrollo Sostenible (Colombia) and handled in accordance with government guidelines on the ethical treatment of animals and all applicable regulations (Estatuto Nacional de Proteccion de los Animales, Ley 84, December 27, 1989) and following the considerations of the Herpetological Animal Care and Use Committee (HACC, 2004).

Other methods are detailed in *Supporting Informaton*.

ACKNOWLEDGMENTS. We thank Adriana Alberti, Corinne Da Silva, and Jean Weissenbach (Genoscope) for *Mabuya* transcriptome sequencing and assembly; Jean Weissenbach for the gift of the GeneBridge4 RH clones; Aaron Bauer and Breda Zimkus for providing Scincidae samples; Olivia Bawa (Gustave Roussy Preclinical Evaluation Platform) for contributions to the histological analyses; Mélanie Polrot (Gustave Roussy Animal Care Facilities) for assistance in mouse handling; David Ribet for help in GeneBridge4 panel data analysis; Cécile Lemaitre, Jhen Tsang, and Marie Dewannieux for help with the MPZL1 phosphorylation assay; and Christian Lavialle for discussion and critical reading of the manuscript. This work was supported by the CNRS, by a grant from the Ligue Nationale contre Le Cancer (Equipe Labellisée) (to T.H.), and by Agence Nationale de la Recherche Grant ANR RETRO-PLACENTA (to T.H.).

1. Lavialle C, et al. (2013) Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. *Philos Trans R Soc Lond B Biol Sci* 368:20120507.
2. Denner J (2016) Expression and function of endogenous retroviruses in the placenta. *APMIS* 124:31–43.
3. Dupressoir A, et al. (2009) Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proc Natl Acad Sci USA* 106:12127–12132.
4. Dupressoir A, et al. (2011) A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. *Proc Natl Acad Sci USA* 108:E1164–E1173.
5. Mi S, et al. (2000) Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403:785–789.
6. Blond JL, et al. (2000) An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *J Virol* 74:3321–3329.
7. Blaise S, de Parseval N, Bénit L, Heidmann T (2003) Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci USA* 100:13013–13018.
8. Dupressoir A, et al. (2005) Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc Natl Acad Sci USA* 102:725–730.
9. Heidmann O, Vernochet C, Dupressoir A, Heidmann T (2009) Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: A new "syncytin" in a third order of mammals. *Retrovirology* 6:107.
10. Cornelis G, et al. (2012) Ancestral capture of syncytin-Car1, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. *Proc Natl Acad Sci USA* 109:E432–E441.
11. Cornelis G, et al. (2013) Captured retroviral envelope syncytin gene associated with the unique placental structure of higher ruminants. *Proc Natl Acad Sci USA* 110: E828–E837.
12. Cornelis G, et al. (2014) Retroviral envelope syncytin capture in an ancestrally diverged mammalian clade for placentation in the primitive Afrotherian tenrecs. *Proc Natl Acad Sci USA* 111:E4332–E4341.
13. Redelsperger F, et al. (2014) Capture of syncytin-Mar1, a fusogenic endogenous retroviral envelope gene involved in placentation in the Rodentia squirrel-related clade. *J Virol* 88:7915–7928.
14. Vernochet C, et al. (2014) The captured retroviral envelope syncytin-A and syncytin-B genes are conserved in the Spalacidae together with hemotrichorial placentation. *Biol Reprod* 91:148.
15. Cornelis G, et al. (2015) Retroviral envelope gene captures and syncytin exaptation for placentation in marsupials. *Proc Natl Acad Sci USA* 112:E487–E496.
16. Blackburn DG (2015) Evolution of vertebrate viviparity and specializations for fetal nutrition: A quantitative and qualitative analysis. *J Morphol* 276:961–990.
17. Roberts RM, Green JA, Schulz LC (2016) The evolution of the placenta. *Reproduction* 152:R179–R189.
18. Blackburn DG, Vitt LJ, Beuchat CA (1984) Eutherian-like reproductive specializations in a viviparous reptile. *Proc Natl Acad Sci USA* 81:4860–4863.
19. Jerez A, Ramírez-Pinilla MP (2001) The allantoplacenta of Mabuya mabouya (Sauria, Scincidae). *J Morphol* 249:132–146.
20. Blackburn DG, Vitt LJ (2002) Specializations of the chorioallantoic placenta in the Brazilian scincid lizard, Mabuya heathi: A new placental morphotype for reptiles. *J Morphol* 254:121–131.
21. Blackburn DG, Flemming AF (2010) Reproductive specializations in a viviparous African skink and its implications for evolution and conservation. *Herpetol Conserv Biol* 5:263–270.
22. Flemming AF, Blackburn DG (2003) Evolution of placental specializations in viviparous African and South American lizards. *J Exp Zool A Comp Exp Biol* 299:33–47.
23. 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. *J Comp Physiol B* 181:575–594.
24. Jerez A, Ramírez-Pinilla MP (2003) Morphogenesis of extraembryonic membranes and placentation in Mabuya mabouya (Squamata, Scincidae). *J Morphol* 258:158–178.
25. Ramírez-Pinilla MP, De Pérez G, Carreño-Escobar JF (2006) Allantoplacental ultrastructure of an Andean population of Mabuya (Squamata, Scincidae). *J Morphol* 267: 1227–1247.
26. Leal F, Ramírez-Pinilla MP (2008) Morphological variation in the allantoplacenta within the genus Mabuya (Squamata: Scincidae). *Anat Rec (Hoboken)* 291:1124–1139.
27. Wooding FBP, Ramirez-Pinilla MP, Forhead AS (2010) Functional studies of the placenta of the lizard Mabuya sp. (Scincidae) using immunocytochemistry. *Placenta* 31: 675–685.
28. Pinto-Sánchez NR, Calderón-Espinosa ML, Miralles A, Crawford AJ, Ramírez-Pinilla MP (2015) Molecular phylogenetics and biogeography of the Neotropical skink genus Mabuya Fitzinger (Squamata: Scincidae) with emphasis on Colombian populations. *Mol Phylogenet Evol* 93:188–211.
29. Zhao R, Guerrah A, Tang H, Zhao ZJ (2002) Cell surface glycoprotein PZR is a major mediator of concanavalin A-induced cell signaling. *J Biol Chem* 277:7882–7888.
30. Zannettino AC, et al. (2003) Novel mesenchymal and haematopoietic cell isoforms of the SHP-2 docking receptor, PZR: Identification, molecular cloning and effects on cell migration. *Biochem J* 370:537–549.
31. Roubelakis MG, Martin-Rendon E, Tsaknakis G, Stavropoulos A, Watt SM (2007) The murine ortholog of the SHP-2 binding molecule, PZR accelerates cell migration on fibronectin and is expressed in early embryo formation. *J Cell Biochem* 102:955–969.
32. Jia D, et al. (2014) Amplification of MPZL1/PZR promotes tumor cell migration through Src-mediated phosphorylation of cortactin in hepatocellular carcinoma. *Cell Res* 24:204–217.
33. Dufaure J, Hubert J (1961) Table de développement du lézard vivipare: Lacerta (Zootoca) vivipara Jacquin. *Arch Anat Microsc Morphol Exp* 50:309–328.
34. Wagner GP, Kin K, Lynch VJ (2013) A model based criterion for gene expression calls using RNA-seq data. *Theory Biosci* 132:159–164.
35. Henzy JE, Johnson WE (2013) Pushing the endogenous envelope. *Philos Trans R Soc Lond B Biol Sci* 368:20120506.
36. Pereira AG, Schrago CG (2017) Arrival and diversification of mabuyine skinks (Squamata: Scincidae) in the Neotropics based on a fossil-calibrated timetree. *PeerJ* 5: e3194.
37. Karin BR, Metallinou M, Weinell JL, Jackman TR, Bauer AM (2016) Resolving the higher-order phylogenetic relationships of the circumtropical Mabuya group (Squamata: Scincidae): An out-of-Asia diversification. *Mol Phylogenet Evol* 102:220–232.
38. Vieira S, de Perez G, Ramírez-Pinilla MP (2007) Invasive cells in the placentome of Andean populations of Mabuya: An endotheliochorial contribution to the placenta? *Anat Rec (Hoboken)* 290:1508–1518.
39. Gyapay G, et al. (1996) A radiation hybrid map of the human genome. *Hum Mol Genet* 5:339–346.
40. Esnault C, et al. (2008) A placenta-specific receptor for the fusogenic, endogenous retrovirus-derived, human syncytin-2. *Proc Natl Acad Sci USA* 105:17532–17537.
41. Boehnke M, Lange K, Cox DR (1991) Statistical methods for multipoint radiation hybrid mapping. *Am J Hum Genet* 49:1174–1188.
42. Zhao ZJ, Zhao R (1998) Purification and cloning of PZR, a binding protein and putative physiological substrate of tyrosine phosphatase SHP-2. *J Biol Chem* 273:29367–29372.
43. Eminaga S, Bennett AM (2008) Noonan syndrome-associated SHP-2/Ptpn11 mutants enhance SIRPalpha and PZR tyrosyl phosphorylation and promote adhesion-mediated ERK activation. *J Biol Chem* 283:15328–15338.
44. Kusano K, Thomas TN, Fujiwara K (2008) Phosphorylation and localization of protein-zero related (PZR) in cultured endothelial cells. *Endothelium* 15:127–136.
45. Ostrovsky AN, et al. (2016) Matrotrophy and placentation in invertebrates: A new paradigm. *Biol Rev Camb Philos Soc* 91:673–711.
46. Blackburn DG, Flemming AF (2012) Invasive implantation and intimate placental associations in a placentotrophic African lizard, Trachylepis ivensi (Scincidae). *J Morphol* 273:137–159.
47. Flemming AF, Branch WR (2001) Extraordinary case of matrotrophy in the African skink Eumecia anchietae. *J Morphol* 247:264–287.
48. Long JA, Trinajstic K, Young GC, Senden T (2008) Live birth in the Devonian period. *Nature* 453:650–652.
49. Wooding P, Burton G (2008) *Comparative Placentation - Structures, Functions and Evolution* (Springer, Berlin).
50. Aiewsakun P, Katzourakis A (2017) Marine origin of retroviruses in the early Palaeozoic Era. *Nat Commun* 8:13954.
51. Henzy JE, Gifford RJ, Kenaley CP, Johnson WE (2017) An intact retroviral gene conserved in spiny-rayed fishes for over 100 My. *Mol Biol Evol* 34:634–639.
52. Metallinou M, et al. (2016) A single origin of extreme matrotrophy in African mabuyine skinks. *Biol Lett* 12:20160430.
53. Ramírez-Pinilla MP, Rueda ED, Stashenko E (2011) Transplacental nutrient transfer during gestation in the Andean lizard Mabuya sp. (Squamata, Scincidae). *J Comp Physiol B* 181:249–268.
54. Gilbert DG (2013) Gene-omes built from mRNA-seq not genome DNA. *Seventh Annu Arthropod Genomics Symposium*, 10.7490/f1000research.1112594.1.
55. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
56. Roberts A, Pachter L (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat Methods* 10:71–73.
57. Hou Z-C, et al. (2012) Elephant transcriptome provides insights into the evolution of eutherian placentation. *Genome Biol Evol* 4:713–725.
58. 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 Biol Evol* 4: 394–411.
59. Lynch VJ, et al. (2015) Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. *Cell Rep* 10:551–561.
60. Eisenberg E, Levanon EY (2013) Human housekeeping genes, revisited. *Trends Genet* 29:569–574.
61. Larsson A (2014) AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278.
62. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59: 307–321.
63. Lefort V, Longueville J-E, Gascuel O (2017) SMS: Smart model selection in PhyML. *Mol Biol Evol* 34:2422–2424.
64. Vidal N, Hedges SB (2005) The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. *C R Biol* 328: 1000–1008.
65. Kumazawa Y (2007) Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. *Gene* 388:19–26.
66. Henzy JE, Coffin JM (2013) Betaretroviral envelope subunits are noncovalently associated and restricted to the mammalian class. *J Virol* 87:1937–1946.
67. de Parseval N, Lazar V, Casella J-F, Benit L, Heidmann T (2003) Survey of human genes of retroviral origin: Identification and transcriptome of the genes with coding capacity for complete envelope proteins. *J Virol* 77:10414–10422.
68. Perro M, et al. (2010) Generation of multi-functional antigen-specific human T-cells by lentiviral TCR gene transfer. *Gene Ther* 17:721–732.