SUPPLEMENTARY MATERIALS

for the article

Contrasted evolution of the vomeronasal receptor repertoires
in Mammals and Squamate reptiles

Urszula Brykczyńska¹*, Athanasia C. Tzika¹*, Ivan Rodriguez², & Michel C. Milinkovitch¹

* These authors contributed equally

¹Laboratory of Artificial & Natural Evolution (LANE), Dept. of Genetics & Evolution, University of Geneva, Sciences III, 30 Quai Ernest-Ansermet, 1211 Geneva, Switzerland;

²Laboratory of Neurogenetics & National Research Center Frontiers in Genetics, Dept. of Genetics & Evolution, University of Geneva, Sciences III, 30 Quai Ernest-Ansermet, 1211 Geneva, Switzerland.

urszula.brykczyńska@yahoo.com, athanasia.tzika@unige.ch, ivan.rodriguez@unige.ch, michel.milinkovitch@unige.ch

Corresponding author:

Michel C. Milinkovitch
Laboratory of Artificial & Natural Evolution (LANE), Dept. of Genetics & Evolution, University of Geneva, Sciences III, 30, Quai Ernest-Ansermet, 1211 Genève 4, Switzerland.
Tel: +41(0)22 379 67 85; Fax: +41(0)22 379 67 95
URL: www.lanevol.org
E-mail: Michel.Milinkovitch@unige.ch
Figure S1. Sequence data assembly scheme.

Figure S2. Validation of contigs. Nucleotide differences distribution function computed for 23 validated contigs, as the number of differences between the nucleotide sequences of the contig and of the PCR product, divided by the product length.
Figure S3. Multiple alignment of 66 corn-snake V2R contigs spanning the 7-trans-membrane region.
**Figure S4.** Evolutionary history of snake and other vertebrate V2Rs. Maximum likelihood tree, based on a multiple alignment of 7TM-domain aa-sequences from 66 corn snake and 15 Python V2Rs (dashed purple branches), as well as 66 representative V2R sequences from five vertebrate species. RaxML branch support values are indicated for major clades. Pink boxes highlight the same five snake-specific expansions than in Figure 3. Arrows indicate sequences used for in-situ hybridization (Fig. 4 and 5), as well as the Python V2R family C sequence. The tree is rooted with taste receptors (TAS1Rs). Scale bar: mean number of aa substitutions per site.
**Figure S5.** Percentages of corn snake VNO sensory neurons expressing various subfamily ABD (left) and subfamily C (right) V2R receptors, analyzed by *in-situ* hybridization. Each data point represents a cell count from one coronal section. Individuals 1 and 5 were juvenile animals.

| Animal ID | Sex   | Age     | State   | Pantherophis V1r1/2 | Pantherophis V1r3/4 |
|-----------|-------|---------|---------|---------------------|---------------------|
| EG024     | female| 4 years | active  | -                   | +                   |
| EG110     | male  | 2.5 years | hibernation | -                   | -                   |
| EG161     | female| 2 years | active  | -                   | +++                  |
| EG258     | male* | 1 year  | active  | -                   | ++                   |
| EG000     | male  | 1.5 years | hibernation | -                   | -                   |

**Figure S6.** VNO expression of Pantherophis V1r genes. (A) In situ hybridization (green) of a coronal section of a male corn snake VNO, with an antisense RNA probe for *Pantherophis* V1r3. DNA is stained with Hoechst (blue). VNE, vomeronasal epithelium; L, lumen. Scale bar, 20 µm. (B) Levels of expression of *Pantherophis* V1r genes in the VNO of different individuals. The image in (A) corresponds to the animal indicated with an asterisk.
Table S1. RNA-seq number of reads and of contig in pre-assembly and final assembly; F1-F4 = Females 1-4, M1-M4 = Males 1-4.

| Assembly step | Sample | Platform | Read length [bp] | Software            | Total reads | Assembled reads / % total | Contigs |
|---------------|--------|----------|------------------|---------------------|-------------|---------------------------|---------|
| Pre-assembly  | F1F2F3, F4, M1M2M3 | Roche 454 | 300-550          | SeqMan              | 343,062     | 270,972 / 79%             | 38,039  |
| Pre-assembly  | F1F2F3 | Illumina | 114              | Velvet/Oases        | 29,161,853  | 17,655,782 / 60.5%        | 60,270  |
| Pre-assembly  | M4     | Illumina | 114              | Velvet/Oases        | 25,233,055  | 14,098,765 / 55.8%        | 110,365 |
| Final assembly | F1F2F3F4M1M2M3M4 | Roche 454 | 300-550          | SeqMan              | 8,048 54,191| 7,550 / 94% 44,138 / 81.5%| 467     |

Table S2. Corrected pairwise distance matrix for snake V1Rs sequences at the nucleotide (lower diagonal - Jukes-Cantor correction - 563 nt) and the aa (upper diagonal - Poisson correction - 187 aa) levels. The distances among Pantherophis V1r1 to 4 sequences are indicated in yellow. The distances between Pantherophis and Python orthologs are highlighted in green.

|               | Panth V1r3 | Panth V1r4 | Petromyzon 10105 | Python V1r2 | Xenopus s33483 | Anolis 25768 | Xenopus s31104 | Xenopus s30158 | Petromyzon 10425 | Panth V1r1 | Panth V1r2 | Xenopus V1R18 | Python V1r1 |
|---------------|------------|------------|-------------------|-------------|----------------|--------------|----------------|----------------|----------------|------------|------------|----------------|-------------|
| Panth V1r3    |            | 0.23       | 1.72              | 0.28        | 2.27           | 0.44         | 0.82           | 2.02           | 1.78           | 2.43       | 2.37       | 2.22           | 2.02        |
| Panth V1r4    | 0.22       |            | 1.69              | 0.33        | 2.27           | 0.46         | 0.86           | 2.02           | 1.71           | 2.43       | 2.37       | 2.11           | 2.02        |
| Petromyzon 10105 | 1.36 | 1.30       | 1.90              | 2.91        | 1.82           | 1.82         | 1.82           | 1.89           | 2.14           | 2.14       | 1.75       | 2.06           |             |
| Python V1r2   | 0.28       | 0.30       | 1.45              | 2.27        | 0.41           | 0.77         | 2.07           | 1.82           | 2.43           | 2.37       | 2.16       | 2.02           |             |
| Xenopus s33483 | 2.06       | 2.15       | 2.22              | 2.13        | 2.33           | 2.39         | 2.61           | 2.16           | 3.15           | 3.03       | 2.39       | 2.90           |             |
| Anolis 25768  | 0.45       | 0.45       | 1.30              | 0.43        | 2.12           | 0.78         | 2.02           | 1.89           | 2.58           | 2.51       | 2.16       | 2.16           |             |
| Xenopus s31104 | 0.68 | 0.72       | 1.22              | 0.67        | 1.96           | 0.65         | 2.12           | 1.78           | 2.51           | 2.59       | 1.86       | 2.16           |             |
| Xenopus s30158 | 1.75 | 1.89       | 1.49              | 1.89        | 1.94           | 1.77         | 1.79           | 1.93           | 1.55           | 1.53       | 1.22       | 1.22           |             |
| Petromyzon 10425 | 1.29 | 1.25       | 1.58              | 1.31        | 1.63           | 1.35         | 1.36           | 1.64           | 2.36           | 2.31       | 1.97       | 2.02           |             |
| Panth V1r1    | 1.74       | 1.63       | 1.62              | 1.79        | 2.22           | 1.68         | 1.77           | 1.14           | 1.65           | 0.23       | 1.12       | 0.54           |             |
| Panth V1r2    | 1.78       | 1.67       | 1.58              | 1.78        | 2.23           | 1.74         | 1.82           | 1.16           | 1.61           | 0.17       | 1.10       | 0.54           |             |
| Xenopus V1R18 | 1.41       | 1.47       | 1.31              | 1.43        | 2.09           | 1.36         | 1.38           | 0.96           | 1.30           | 0.90       | 0.90       | 0.87           |             |
| Python V1r1   | 1.60       | 1.57       | 1.63              | 1.61        | 2.19           | 1.58         | 1.68           | 1.09           | 1.38           | 0.30       | 0.30       | 0.76           |             |
Table S3. Primer sequences used for validation of contigs. Primers used for in situ probe preparation are indicated in grey.

| Primer name       | Sequence                     | Primer name       | Sequence                     |
|-------------------|------------------------------|-------------------|------------------------------|
| V2R-EG001_F       | CCAACTGGTGACCAGAAGGGGA      | V2R-EG112_F       | ACAGTGCTGGAGATGAAGTCTCCT     |
| V2R-EG001_R       | CCCCTGTGACTCACCACCA         | V2R-EG112_R       | CCAAATCACAGTGCCAATCTCA       |
| V2R-EG012_F       | TGACCAAGATCGCTTTCAAGGC     | V2R-EG154_F       | TGACCCCATGGAATCTCAGACAGCA    |
| V2R-EG012_R       | AGATGCAAGCAGAGGAGCCA       | V2R-EG154_R       | CCCCCTGGCTCAAGTAGCTGGG       |
| V2R-EG033_F       | AGGAACAGATGACTATGATCCCCCA  | V2R-EG156_F       | GCTGTATAGATGTCCTGTGCTG       |
| V2R-EG033_R       | TGCTCTTTGTGTCTAGCCTGTC     | V2R-EG156_R       | GAAGACTGTCAGGCGCCACACCA      |
| V2R-EG041_F       | GGCAGATGCATCATCAATGGCCA    | V2R-EG199_F       | AGGGAGGCTGTGCTCTCAAGA       |
| V2R-EG041_R       | CTGTTGAGTGGTCTGAGACAGC     | V2R-EG199_R       | GCAGAGCTGGCGCCAGGTTTCC       |
| V2R-EG042_R       | CCCAGGCCAGAGGCGCAGGAG     | V2R-EG200_F       | TCACAGGCTCTCCTCCACATGGA      |
| V2R-EG044_F       | TGGTTTCCAGTTCCAGAGACAGT    | V2R-EG200_R       | TCAGAGGTGATGGCGCCAGGAT       |
| V2R-EG044_R       | TCATTTCTGCTTATTCTCCAGAG    | V2R-EG204_F       | AGGGAGGCTGTGCTCTCAAGA       |
| V2R-EG049_F       | TGGAATGCAATGCGCTTTGCTG    | V2R-EG222_F       | GATGATGCTCAAGAAATGCTTGGCT    |
| V2R-EG049_R       | GCAACCAACAGTCCAGACACCA     | V2R-EG223_F       | GCAACCAACAGGATGCGCCAGGA      |
| V2R-EG049_R       | AGGACACCCAGACTCGAGAGC     | V2R-EG233_F       | GCAACCAACAGGATGCGCCAGGA      |
| V2R-EG057_F       | TGGGAAGGTGGCAGTCTGTGTTG    | V2R-EG235_F       | GACACACACAGAAGATGGGACCT      |
| V2R-EG057_R       | AGGTCAAGGCTTCTGTTGCTG     | V2R-EG235_R       | TGCTGGATGTCGGCTCTCAAGC      |
| V2R-EG059_F       | CCACGTCTCCGTGGTGAAGACCA    | EG_OR1_F1         | TTGTGACCTCCTGGTGAGGACATGG   |
| V2R-EG059_R       | TGGGGTGAAGTTCTAACAGCCAT    | EG_OR1_R1         | ACTCGGCGGTAGAAATGCTGAGCTG    |
| V2R-EG068_F       | TGCGATTCCAAGAGGACAGT      | EG_V1ra_F1a       | GTGATGAGCAGGCTCCAGGATCCTC    |
| V2R-EG068_R       | TGAAGGAGCATCCTACAGCCACATG | EG_V1ra_R1a       | ATCCAGAGATGTCGTACTACCC       |
| V2R-EG070_F       | TGGTGTTCTGCTGCTGGCTG      | EG_V1rb_F2        | AGGGATGATGCTATGCTCTTGCT      |
| V2R-EG070_R       | TGATGGTCTGCTGCTGGCTG      | EG_V1rb_R2a       | CCATGATAGGATGAGATCACAGA      |
| V2R-EG107_F       | ACAAGGAAGCTGTGGGCTGCTCTCA | EG_V1rb_R4a       | GACTTTCTGGCTCACACCAACCA      |
| V2R-EG107_R       | AGCCACACACAGAAATGCA       |                   |                              |
| V2R-EG108_F       | TCAACTGGTTTCCCTCTCCCAATCA |                   |                              |
| V2R-EG108_R       | AGGAGGCTGTGCTCTCAAGA      |                   |                              |

Supplementary files available at [www.reptilian-transcriptomes.org](http://www.reptilian-transcriptomes.org)

- **Suppl. File 1.** Sequences of putative functional V2R transcript fragments.
- **Suppl. File 2.** Sequences of putatively-expressed V2R pseudogene fragments.
- **Suppl. File 3.** List of putative Python and Garter snake V2Rs along with the best TBLASTX hit obtained against the NCBI nr database (May 2012).
- **Suppl. File 4.** ORF sequences of Python putative V2Rs.
- **Suppl. File 5.** Corn snake putative olfactory genes.
- **Suppl. File 6a.** Assembled Illumina contigs
- **Suppl. File 6b.** Assembled (contigs) and unassembled (singletons >200bp) Roche/454 sequences.
- **Suppl. File 7.** Sequences of putative reptilian V1Rs.