Evolutionary genetics of pheromonal communication in mouse lemurs

INAUGURAL – DISSERTATION
zur Erlangung des Grades eines Doktors der Naturwissenschaften
– Doctor rerum naturalium –
(Dr. rer. nat.)

vorgelegt von
Philipp Hohenbrink geb. Kappel
Schwerin, Deutschland

Hannover 2014
Wissenschaftliche Betreuung

Prof. Dr. Ute Radespiel, Institut für Zoologie, Tierärztliche Hochschule Hannover

und

Dr. Nicholas I. Mundy, Department of Zoology, University of Cambridge, Vereinigtes Königreich

1. Gutachterin / Gutachter: Prof. Dr. Ute Radespiel, Institut für Zoologie, Tierärztliche Hochschule Hannover

und

Dr. Nicholas I. Mundy, Department of Zoology, University of Cambridge, Vereinigtes Königreich

2. Gutachterin / Gutachter: Prof. Dr. Miguel Vences, Zoologisches Institut, Technische Universität Braunschweig

Tag der mündlichen Prüfung: 19. Mai 2014

Diese Dissertation wurde finanziert von der VolkswagenStiftung.
Index of contents

Index of contents.................................................................I
Previously published or submitted excerpts of this thesis .....................III
List of abbreviations..................................................................V
Summary ..................................................................................VII
German summary (Zusammenfassung) ............................................IX
1. General introduction..............................................................1
  1.1 Olfaction.............................................................................1
  1.2 Terminology: pheromones, signature mixtures and more ..........1
  1.3 Olfactory system in mammals ............................................5
  1.4 Evolutionary origin of the vomeronasal organ .......................6
  1.5 Vomeronasal organ in mammals .........................................7
  1.6 Vomeronasal organ in humans ..........................................8
  1.7 Olfactory and vomeronasal receptors ....................................9
  1.8 Vomeronasal organ and vomeronasal receptor repertoire in primates .................................................13
  1.9 Model system: mouse lemur .............................................15
  1.10 Aims & Hypotheses .....................................................17
  1.11 References ...................................................................18
2. Pervasive and ongoing positive selection in the vomeronasal-1 receptor (V1R) repertoire of mouse lemurs...........................................27
  2.1 Introduction ....................................................................28
  2.2 Materials and Methods ..................................................31
     Data collection .....................................................................31
     Data analysis ......................................................................32
  2.3 Results ...........................................................................34
     Selection on the mouse lemur V1R repertoire ........................ 34
     Distribution of sites under positive selection across the V1R protein .....................................................36
     Selection on individual loci ..................................................38
     Comparison with mouse V1R clades and potential function of mouse lemur V1R gene clusters .............39
     Comparison with another strepsirrhine ................................ 39
  2.4 Discussion ......................................................................40
     Patterns of selection pressures on V1R genes .........................40
     Selection on protein level ....................................................42
  2.5 References ....................................................................43
  2.6 Appendix ......................................................................46
3. First evidence for functional vomeronasal 2 receptor genes in primates ..........................................................51
  3.1 Introduction ....................................................................51
  3.2 Material and methods ....................................................52
     Data collection .....................................................................52
     Confirmation of transcription .............................................53
     Mode of selection on V2R genes .........................................53
  3.3 Results and discussion ....................................................53
     V2Rs in strepsirrhines and tarsiers .......................................53
     Degeneration of VN2R1 and VN2R2 in anthropoid primates ..........55
     Functional significance of V2Rs in strepsirrhines ......................56
  3.4 References ....................................................................57
4. Functional promiscuity in a mammalian chemosensory system: Extensive expression of vomeronasal receptors in the main olfactory epithelium of mouse lemurs .......... 61
   4.1 Introduction ........................................................................................................... 62
   4.2 Materials and Methods .......................................................................................... 63
     Tissue collection ....................................................................................................... 63
     Confirmation of transcribed loci .............................................................................. 65
   4.3 Results .................................................................................................................... 66
   4.4 Discussion ............................................................................................................... 69
   4.5 References ............................................................................................................. 73

5. Signals of a divergent evolutionary history in the vomeronasal receptor genes of two sympatric mouse lemur species ................................................................. 77
   5.1 Introduction ............................................................................................................. 78
   5.2 Material and Methods ........................................................................................... 80
     Data collection ......................................................................................................... 80
     Data analyses .......................................................................................................... 81
   5.3 Results .................................................................................................................... 83
     Differences in genetic diversity ............................................................................... 83
     Divergent demographic histories .......................................................................... 85
     Evidence for purifying selection ............................................................................ 87
   5.4 Discussion ............................................................................................................... 88
     Demographic histories of two sympatric mouse lemur species ............................... 88
     Purifying selection in the recent evolutionary history of vomeronasal receptor genes ... 90
   5.5 References ............................................................................................................. 91

6. General discussion ...................................................................................................... 99
   6.1 Clusters and families ............................................................................................. 99
   6.2 The repertoires of vomeronasal receptors in retrospective ...................................... 102
   6.3 Selection pressures on vomeronasal receptors: a revision .................................... 106
   6.4 Overall conclusion ............................................................................................... 108
   6.5 References ............................................................................................................. 109

Statement (Erklärung) .................................................................................................... 113
Acknowledgements (Danksagung) ................................................................................ 115
Previously published or submitted excerpts of this thesis

Chapters 2, 3 and 4 of this thesis are published in scientific peer-reviewed journals or have been submitted for publication:

Chapter 2 has been published in December 2012 in the journal “Molecular Biology and Evolution” by Philipp Hohenbrink, Ute Radespiel and Nicholas I. Mundy under the title “Pervasive and ongoing positive selection in the vomeronasal-1 receptor (V1R) repertoire of mouse lemurs” (29:3807-3816; doi: 10.1093/molbev/mss188). It was first published online on July 20th 2012. The copyright is held by Oxford University Press since July 20th 2012. Two license agreements between the author of this thesis, Philipp Hohenbrink, and Oxford University Press allow PH to use the text (License Number 3343651028182), figures and tables (License Number 3343651298364) of the publication. Molecular Biology and Evolution is the original place of publication which can be accessed on http://mbe.oxfordjournals.org/content/early/2012/07/19/molbev.mss188. The format of the chapter was adapted to the style of this thesis.

The conception and design of the study presented in chapter 2 was done by PH, UR and NM. The data were acquired and analysed by PH. The interpretation of the data was done by PH, UR and NM. The article was drafted by PH and critically revised by all authors. All authors approved the publication of the final version.

Chapter 3 has been published in February 2013 in the journal “Biology Letter” by Philipp Hohenbrink, Nicholas I. Mundy, Elke Zimmermann und Ute Radespiel under the title “First evidence for functional vomeronasal 2 receptors in primates” (9:20121006). It was first published online on December 26th 2012. The author, PH, retained the copyright in the article. Since February 23rd 2014 PH is allowed to use the article in any “Author Generated Postprint” including this thesis.

The conception and design of the study presented in chapter 3 was done by PH, NM and UR. Material was provided by EZ. The data were acquired and analysed by PH. The interpretation of the data was done by PH, NM and UR. The article was drafted by PH and critically revised by all authors. All authors approved the publication of the final version.

Chapter 4 has been submitted to the journal “Chemical Senses” by Philipp Hohenbrink, Silke Dempewolf, Elke Zimmermann, Nicholas I. Mundy and Ute Radespiel
under the title “Functional promiscuity in a mammalian chemosensory system: Extensive expression of vomeronasal receptors in the main olfactory epithelium of mouse lemurs” for publication in its current form. NM and UR contributed equally to the work.

The conception and design of the study was done by PH, NM and UR. Material was provided by EZ. The data were acquired by PH and SD and analysed by PH. The interpretation of the data was done by PH, NM and UR. The article was drafted by PH and critically revised by all authors. All authors approved the publication of the final version.

Chapter 5 is intended to be submitted by Philipp Hohenbrink, Nicholas I. Mundy and Ute Radespiel to a scientific peer-reviewed journal for publication in its current or slightly revised form.

The conception and design of the study was done by PH, NM and UR. The data were acquired and analysed by PH. The interpretation of the data was done by PH, NM and UR. The article was drafted by PH and critically revised by all authors.
**List of abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>beta-actin</td>
</tr>
<tr>
<td>AOB</td>
<td>Accessory olfactory bulb</td>
</tr>
<tr>
<td>App.</td>
<td>Appendix</td>
</tr>
<tr>
<td>BEB</td>
<td>Bayes Empirical Bayes</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair(s)</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>cDNA</td>
<td>copy DNA (Deoxyribonucleic acid)</td>
</tr>
<tr>
<td>D-loop</td>
<td>Displacement loop</td>
</tr>
<tr>
<td>df</td>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>$d_N$</td>
<td>Number of nonsynonymous substitutions per site</td>
</tr>
<tr>
<td>$d_S$</td>
<td>Number of synonymous substitutions per site</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>e.g.</td>
<td>exempli gratia (“for example”)</td>
</tr>
<tr>
<td>ESP1</td>
<td>Exocrine gland-secreting peptide 1</td>
</tr>
<tr>
<td>et al.</td>
<td>et alii/aliae (“and other [coauthors]”)</td>
</tr>
<tr>
<td>Fig.</td>
<td>Figure</td>
</tr>
<tr>
<td>FPR</td>
<td>Formyl peptide receptor</td>
</tr>
<tr>
<td>G-protein</td>
<td>GTP (Guanosine triphosphate)-binding protein</td>
</tr>
<tr>
<td>GC-D</td>
<td>Guanylyl-cyclase-D</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptors</td>
</tr>
<tr>
<td>i.e.</td>
<td>id est (“that is”)</td>
</tr>
<tr>
<td>JBA</td>
<td>Jardin Botanique A</td>
</tr>
<tr>
<td>LRS</td>
<td>Likelihood ratio statistics</td>
</tr>
<tr>
<td>LRT</td>
<td>Likelihood ratio test</td>
</tr>
<tr>
<td>mGluR</td>
<td>Metabotropic glutamate receptor</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MKT</td>
<td>McDonald-Kreitman test</td>
</tr>
<tr>
<td>MOB</td>
<td>Main olfactory bulb</td>
</tr>
<tr>
<td>MOE</td>
<td>Main olfactory epithelium</td>
</tr>
<tr>
<td>MUP</td>
<td>Major urinary protein</td>
</tr>
<tr>
<td>OR</td>
<td>Olfactory receptor</td>
</tr>
<tr>
<td>PAML</td>
<td>Phylogenetic analysis by Maximum Likelihood</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>TRPC2</td>
<td>Transient receptor potential channel type C2</td>
</tr>
<tr>
<td>VNO</td>
<td>Vomeronasal organ</td>
</tr>
<tr>
<td>V1R</td>
<td>Vomeronasal receptor type 1</td>
</tr>
<tr>
<td>V2R</td>
<td>Vomeronasal receptor type 2</td>
</tr>
<tr>
<td>VR</td>
<td>Vomeronasal receptor</td>
</tr>
<tr>
<td>WGA</td>
<td>Whole genome amplification</td>
</tr>
</tbody>
</table>
Summary
Philipp Hohenbrink
Evolutionary genetics of pheromonal communication in mouse lemurs

This thesis investigated the evolutionary genetics of the vomeronasal system of mouse lemurs, nocturnal primates which rely heavily on olfactory signalling. The vomeronasal system is one of two systems for olfactory perception in mammals and is mainly used for pheromonal intraspecific communication. Most previous studies focused on rodents, where two types of vomeronasal receptor (VR) genes are present in the vomeronasal organ (VNO, located at the base of the nasal cavity): vomeronasal receptor type 1 and type 2 (V1R and V2R). The translated proteins are structurally different, but information about the evolution of VRs in primates is limited. Mouse lemurs have the largest repertoire of V1Rs among primates. However, the presence of functional V2Rs had not been reported. Four studies were designed to answer specific questions and to analyse separate aspects of the evolution of the VR repertoire in mouse lemurs.

The first study investigated the structure of the V1R repertoire in the best studied mouse lemur species, the grey mouse lemur, and identified nine monophyletic gene clusters. It was shown that seven of nine gene clusters evolved under strong positive selection indicating that adaptation of V1Rs influenced the evolution of the species. Comparison with function of related V1R clusters in mice suggested a potential relationship between receptor function and strength of selection. Positive selection was also detected within five of seven analysed V1R loci sequenced from six to ten mouse lemur species, indicating ongoing selection within the genus, which may be related to sexual selection and, potentially, speciation processes.

In the second study it was confirmed for the first time that two intact V2R genes are present in the genomes of three strepsirrhine primates (lemurs and lorisoids) and both were expressed in the VNO of the grey mouse lemur. The conservation of V2R genes contrasted to the highly diverse V1R repertoire and indicated divergent evolution of both types of VRs.

The analysis of expression was expanded in the third study where specific expression patterns of 64 V1Rs and the two V2Rs were investigated in the VNO and in the main olfactory epithelium (MOE, located in the upper nasal cavity) from different mouse lemur individuals. All loci were expressed in the VNO of at least one animal which showed individual proportions of expressed VR loci of 83% to 97%. Unexpectedly, extensive expression of VRs...
was also detected in the MOE, with proportions of expressed loci in individuals ranging from 29% to 45%. *TRPC2*, a cation channel protein crucial for signal transduction via VRs, was coexpressed in both organs indicating likely functionality of expressed VRs also in the MOE. Given the differences in the neural pathways of MOE and VNO signals, which project to higher cortical brain centres or the limbic system, respectively, this raises the possibility that the evolution of MOE-expression of VRs enabled mouse lemurs to adaptively diversify the processing of VR-encoded olfactory information.

The fourth study analysed the genetic variation of VRs within wild populations of two sympatric mouse lemur species and helped to distinguish genes from alleles. In contrast to the results on the species level that are rather reflecting evolutionary processes in the distant past (previous chapters), this study indicated that purifying selection is the prominent mode of selection in the more recent evolutionary past of mouse lemurs at the population level. Additionally, the comparison of the genetic diversity and demographic modelling of 15 *V1R* and the two *V2R* loci revealed signals of a divergent demographic history in the two sympatric species.

In conclusion, the studies showed the complexity of the VR system in the small nocturnal primates that has evolved under varying selection pressures. However, information about specific corresponding ligands is currently lacking in mouse lemur species, which would be needed to combine the genetic results with the actual ecological function of VRs. Furthermore, the genome of the grey mouse lemur is incompletely sequenced and a high-coverage genome is necessary to identify the full repertoire of *V1R* and *V2R* loci. Additional genome data from other lemurs and lorisoids would help to identify VR repertoires among strepsirrhine primates and reveal if the complexity in mouse lemurs is unique or typical for lemurs.
German summary (Zusammenfassung)
Philipp Hohenbrink
Evolutionsgenetik pheromongebundener Kommunikation bei Mausmakis

In dieser Arbeit wurde die Evolutionsgenetik des vomeronasalen Systems bei Mausmakis untersucht, nachtaktiven Primaten, die stark auf olfaktorische Signale angewiesen sind. Das vomeronasale System ist eines von zwei Systemen zur olfaktorischen Wahrnehmung bei Säugetieren und wird hauptsächlich zur pheromongebundenen innerartlichen Kommunikation verwendet. Die meisten bisherigen Studien waren auf Nagetiere fokussiert, bei denen zwei Gen-Typen von Vomeronasalrezeptoren (VR) im Vomeronasalorgan (VNO, am Boden der Nasenhöhle) vorhanden sind: Vomeronasalrezeptor Typ 1 und Typ 2 (V1R und V2R). Die translatierten Proteine beider Typen sind strukturell unterschiedlich, allerdings liegen keine hinreichenden Informationen über die Evolution von VRs bei Primaten vor. Mausmakis haben von allen Primaten das größte Repertoire an V1R-s, wobei die Präsenz von funktionellen V2R-s bis jetzt nicht bestätigt wurde. Es wurden vier Studien ausgearbeitet, um spezifische Fragestellungen zu beantworten und verschiedene Aspekte der Evolution des V1R-Repertoires bei Mausmakis zu untersuchen.

Die erste Studie untersuchte den Aufbau des V1R-Repertoires in der am besten untersuchten Mausmaki-Art, dem grauen Mausmaki, und identifizierte neun monophyletische Gen-Cluster. Es konnte gezeigt werden, dass sieben dieser neun Gen-Cluster unter starker positiver Selektion evolviert sind, was darauf hindeutet, dass Adaption von V1Rs die Evolution dieser Art beeinflusst hat. Vergleiche mit der Funktion von verwandten V1R Clustern in der Maus ließen auf einen potentiellen Zusammenhang zwischen der Funktion des Rezeptors und der Stärke an Selektion schließen. Positive Selektion wurde auch in fünf von sieben analysierten V1R Loci detektiert, die von sechs bis zehn verschiedenen Mausmaki-Arten sequenziert wurden, was auf anhaltende Selektion innerhalb der Gattung hindeutet und in Bezug zu sexueller Selektion oder eventuell auch zu Artbildungsprozessen stehen könnte.

In der zweiten Studie wurde zum ersten Mal gezeigt, dass sich zwei intakte V2R Gene in den Genomen von drei Feuchtnasenaffen (Lemuren und Loris) befinden und beide im VNO des grauen Mausmakis exprimiert werden. Die Konservierung der V2R Gene steht im Kontrast zum hoch diversen V1R-Repertoire und weist auf unterschiedliche Evolution der beiden VR-Typen hin.
Die Analyse der Expression wurde in der dritten Studie ausgeweitet, in der spezifische Expressionsmuster von 64 V1Rs und den zwei V2Rs im VNO und der Riechschleimhaut (MOE, in der oberen Nasenhöhle) in mehreren Mausmaki-Individuen untersucht wurde. Alle Loci wurden im VNO mindestens eines Tieres exprimiert, wobei die individuellen Anteile an exprimierten VR-Loci zwischen 83% und 97% schwankten. Unerwarteterweise wurde auch erhebliche Expression von VRs im MOE festgestellt, wobei hier die individuellen Anteile an exprimierten Loci zwischen 29% und 45% lagen. TRPC2, ein Kationenkanal-Protein, das unverzichtbar für die VR-übermittelte Signalübertragung ist, war in beiden Organen exprimiert, was auf mögliche Funktionalität der exprimierten VRs auch im MOE hinweist. Angesichts der Unterschiede in den Nervenbahnen von MOE und VNO Signalen, die zu höheren kortikalen Gehirnarealen respektive dem Limbischen System führen, ergibt sich die Möglichkeit, dass die Evolution von VR-Expression im MOE Mausmakis ermöglicht hat, die Verarbeitung von VR-vermittelten olfaktorischen Informationen zu diversifizieren.

Die vierte Studie analysierte die genetische Diversität von VRs in wild lebenden Populationen zweier sympatrischer Mausmaki-Arten und half bei der Unterscheidung von Genen und Allelen. Im Gegensatz zu den Ergebnissen auf Artenebene, welche eher die evolutionären Prozesse in der weiter zurückliegenden Vergangenheit wiedergeben (vorherige Kapitel), zeigte diese Studie, dass negative Selektion der treibende Selektionsdruck während der jüngeren Vergangenheit von Mausmakis auf Populationsebene war. Zusätzlich konnten ein Vergleich der genetischen Diversität und eine demographische Modellierung von 15 V1Rs und den zwei V2Rs aufzeigen, dass beide sympatrischen Arten eine unterschiedliche demographische Geschichte vorweisen.

1. General introduction

1.1 Olfaction

Olfaction – or the sense of smell – is an important sensory ability throughout the animal kingdom. It is used to detect food, predators, conspecifics and even environmental dangers such as fire. This introductory part will focus on tetrapods, especially mammals, but basically every animal is capable of olfactory perception and the principles are quite the same in mammals and, for example, in insects. Odour molecules bind to more or less specific receptors and depolarize nerve cells that lead to processing in the brain. Unlike gustation – or taste –, the other chemical sense, olfaction uses sensory nerve cells that directly project into the brain with their axons. The combination of all smell perceiving organs and their subsequent processing areas in the brain (primarily the olfactory bulb) is called the “olfactory system”.

1.2 Terminology: pheromones, signature mixtures and more

A so-called semiochemical is a molecule that is used between organisms to interact chemically, mostly by olfaction. Semiochemicals are distinguished from hormones which are used within an organism (Nordlund and Lewis 1976). This thesis utilizes the most recent differentiation of semiochemicals into pheromones and signature mixtures (used between members of the same species) or allelochemicals (used between members of different species, Wyatt 2014, pp. 2-15). Although the three classes may show differences in composition or molecular weight, the basic difference is their origin. In contrast to allelochemicals, pheromones and signature mixtures are compounds which are actively or passively released by a conspecific and used for communication within the species. Hence the same molecule can be a pheromone in one species but an allelochemical in another species. Allelochemicals are further divided into allomones (benefit the emitter), kairomones (benefit the receiver) or synomones (benefit both emitter and receiver).

Basically all living creatures like animals, plants and even bacteria emit molecules as by-products of their metabolism, during excretion of waste products or on purpose (Touhara and Vosshall 2009). If such molecules can be perceived by the olfactory system of animals, it is called an “odorant”. Odorants are usually volatile and have a molecular weight below 300 Da. The term “odorant” overlaps with the term “semiochemical” as many odorants are semiochemicals and vice versa. However, there are even inorganic odorants like sulphur...
Dioxide and even some metals can be perceived by olfaction (Touhara and Vosshall 2009). In general, selection should favour the perception of odorants that are important for survival whereas the ability to perceive molecules without any informative content should not evolve. Perceived odorants give animals essential information e.g. about the location of food, the ripeness of fruits and the presence of predators or fire.

The word “pheromone” was first used in 1959 by Karlson and Lüscher (1959). It is derived from the Greek words φέρειν (pherein = to transfer) and όρμον (hormon = to excite). A pheromone is a substance or a combination of multiple compounds that is secreted by one individual and received by a conspecific in which it induces specific behaviours (releaser pheromone) or physiological changes (primer pheromone). By definition it has to be transferred between conspecifics, i.e., pheromones are used for intraspecific communication. Pheromones are evolved signals in contrast to most other semiochemicals that did not evolve to convey information (Wyatt 2014, pp. 5-6). As indicated above, a chemical compound secreted by one individual can be a pheromone for a second conspecific but can function as a kairomone for a third individual of another species that preys upon the compound secreting species (Touhara and Vosshall 2009). Furthermore, allelochemicals can also induce specific behaviours or physiological changes similar to pheromones. For example, nectarivorous insects are attracted by kairomones released by flowers that induce feeding behaviour. Nevertheless, the term pheromone is inappropriate in this case as the responsible compound is not released by a conspecific. Pheromones probably derived from molecules that had other functions e.g. hormones, host plant odorants or waste products and then gained a new function as a pheromone in the course of evolution (Wyatt 2003, p. 6). Selection would favour longevity and specificity as well as an enhancement of the signalling function, whereby the original function may be lost. Pheromones are usually actively secreted in urine, sweat or saliva but mammals in particular possess specific scent glands primarily used for pheromone synthesis, storage and release of pheromonal secretions.

The term “pheromone” has been used more loosely over the last decades, e.g., for molecules used in intraspecific communication without inducing obvious behavioural or physiological changes, such as individual recognition. Therefore, Wyatt introduced the term “signature mixtures” in 2010 which are “variable subsets of molecules of an animal’s chemical profile which are learnt by other animals, allowing them to distinguish individuals or colonies” (Wyatt 2010, p. 685). In contrast to the learning process linked to signature mixtures, the classical pheromone induces innate responses. Regardless, those responses can
be modified depending on experience or context. This thesis uses the classical definition of pheromones and distinguishes them from signature mixtures.

Animals can use different types of signals to communicate with conspecifics. When comparing these types, olfactory signals (here equal to pheromones) have advantages and disadvantages over acoustic, visual or tactile signals (Table 1.1). Pheromones can be dispersed over a long distance even in a dense environment with low energetic costs. In return, it is difficult for a receiver to locate the sender of the olfactory signal. Pheromones can also be used for marking as long-lasting chemicals can display a territory during the absence of the sender. This is not possible with acoustic or visual signals as they are instantaneous. Pheromones with varying durability can be used in different situations and have evolved by selection acting on their chemical characteristics such as volatility or stability (Wyatt 2003, p. 13). For example, a study in rodents, carnivores and primates showed that sex pheromones are smaller and therefore more volatile and short-lived than pheromones used for territory marking (Alberts 1992).

Table 1.1: Characteristics of different communication channels (after Alcock 1989, p. 246, Table 3; and Wyatt 2003, p. 12, Table 1.2)

<table>
<thead>
<tr>
<th>Feature of channel</th>
<th>Type of signal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Olfactory</td>
</tr>
<tr>
<td>Range</td>
<td>Long</td>
</tr>
<tr>
<td>Transmission rate</td>
<td>Slow</td>
</tr>
<tr>
<td>Flow round barrier?</td>
<td>Yes</td>
</tr>
<tr>
<td>Locatability of sender</td>
<td>Difficult</td>
</tr>
<tr>
<td>Energetic cost to sender</td>
<td>Low</td>
</tr>
<tr>
<td>Longevity</td>
<td>Variable, potentially high</td>
</tr>
<tr>
<td>Use in darkness</td>
<td>Yes</td>
</tr>
<tr>
<td>Specificity</td>
<td>Potentially very high</td>
</tr>
</tbody>
</table>
The extent of pheromone usage for intraspecific communication may be related to habitat use and daily activity rhythm (Wyatt 2003, p. 13). For example, the grysbok, *Raphicerus melanotis*, a small solitary antelope lives in dense scrub and relies heavily on olfactory communication (Albone 1984, p. 91), more than antelope species in more open habitats (Wyatt 2003, p. 13). Also, nocturnal animals rely more on olfactory or acoustic cues than on visual cues (e.g. Barton et al. 1995; Zimmermann 1995).

Pheromones that are transmitted through air should, in general, have a size of 5 – 20 carbon atoms and a molecular weight of 80 – 300 Da (Wilson and Bossert 1963, p. 685; Wilson 1970, pp. 144-146). Only a few kinds of pheromone molecules below this size range can be generated and stored by glandular tissue. Furthermore, molecular diversity increases rapidly with molecular weight; hence sex pheromones are expected to be in the upper size range of airborne compounds (molecular weight of 200 – 300 Da) to allow more specificity (Wilson and Bossert 1963, p. 685). However, large molecules are energetically more expensive to generate – especially in such huge amounts needed for airborne pheromones – and they tend to be less volatile leading to a decreased dispersability (Wilson 1970, p. 146). On the other hand, pheromones can be non-volatile organic compounds, peptides or proteins with a molecular weight of up to 10,000 Da. Such pheromones should preferably be used in close contact situation. Also molecules used for signature mixtures that inform about individuality should probably be less volatile since they can be more reliably associated with the producer if they do not disperse easily (Brennan and Zufall 2006).

Pheromones can become quite species-specific after being under selective pressure from sexual selection and speciation (Wyatt 2003, p. 16). Basically two ways to gain specificity are common: first, the evolution of large molecules – for example, peptides – and second, the evolution of multi-component pheromones consisting of several more simple molecules that stimulate in combination. A pheromone in mice urine, for example, elicits aggression in other males and consists of two compounds that are both inactive separately (Novotny et al. 1985). On the other hand, some pheromones, particularly alarm pheromones, do not need to be exclusive and are often not species-specific. If a species releases alarm pheromones at the presence of a predator and if it is beneficial for other species to respond, then cross-species responses will evolve (Wyatt 2003, p. 16). Because of the omnipresence of chemical cues used for chemical communication among animals, they are also most likely under sexual selection (Wyatt 2003, p. 22).
1.3 Olfactory system in mammals

In mammals the organs used for olfaction are roughly located in or around the nasal cavity and the mammalian olfactory system basically consists of two subsystems: the main and the accessory olfactory system. Other subsystems are known in mice and – considering that mice are among the best-studied model organisms – undiscovered subsystems in other less studied species are likely. Such subsystems in mice include the septal organ of Masera, the Grueneberg ganglion and the trigeminal nerve (for a review of the subsystem organisation in mammals see Munger et al. 2009).

The main olfactory system consists of the main olfactory epithelium (MOE) and the main olfactory bulb (MOB). The accessory olfactory system consists of the vomeronasal organ (VNO) and the accessory olfactory bulb (AOB, Fig. 1.1). It was previously thought that the main olfactory system is used for odorant detection whereas the accessory olfactory system is exclusively used for pheromone detection (dual olfactory hypothesis, Raisman 1972; Scalia and Winans 1975). Such strong separation was found to be wrong (reviewed by Restrepo et al. 2004; Baxi et al. 2006; Zufall and Leinders-Zufall 2007). Some pheromones are detected in the MOE whereas some odorants without pheromonal information are perceived by the VNO. For example, a pheromone that induces nipple-search behaviour in rabbit pups is perceived by the MOE and not affected by surgical VNO removal (Hudson and Distel 1986; Charra et al. 2012). On the other hand, mice can detect rat predators by perceiving odorants in the rat urine with the VNO (Papes et al. 2010). Interestingly, these odorants are rat MUPs (major urinary proteins) that are closely related to mice MUPs used in pheromonal communication (Hurst et al. 2001).
1.4 Evolutionary origin of the vomeronasal organ

The VNO was described by Ludvig Jacobson in 1813 (English translation from Danish original: Trotier and Døving 1998) and is therefore also known by the name Jacobson’s organ. The VNO is not exclusive to mammals but can be found throughout tetrapod lineages including most amphibians and reptiles, but is absent in birds (Parsons 1967; Stoddard 1980, pp. 17-18). However, the evolution of this organ started in sarcopterygian fish (Bertmar 1981) and though critically discussed in the past (Rudebeck 1944; Parsons 1967) more recent studies described the vomeronasal system (including AOB) in lungfish (González et al. 2010; Nakamuta et al. 2012). Bertmar proposed his theory about the evolution of the VNO in vertebrates based on own earlier studies and literature research (for more references see Bertmar 1981). First used as diverticula (or sacs) from the olfactory organs to improve the olfactory function, the VNO separated from the MOE during evolution. As both olfactory organs originated in aquatic animals they need to produce secretions to function on land. Serous glands have therefore evolved in the MOE and the vomeronasal submucosa (Bertmar 1981). A comparative study showed a reasonably uniform histology and cytology of vomeronasal sensory epithelium and olfactory epithelium in anurans, reptiles and mammals (Graziadei 1971).
The VNO varies immensely in morphology and position among amphibians (Eisthen 2000) and is, for example, involved in reproductive signalling (Wirsig-Wiechmann et al. 2002). It is still more or less a sac of the nasal cavity in all amphibians, but completely detached in amniotes, if present (Bertmar 1981). In most reptiles the VNO is well developed especially in snakes but it is absent in crocodiles and chameleons (Stoddard 1980, pp. 17-18). The nasal duct (which connects the nasal cavity with the VNO) closed during evolution (and closes during embryogenesis) in snakes and ground-living lizards, but the VNO is currently connected with the oral cavity through the palatine duct (Bertmar 1981). In these groups airborne chemicals are collected by flicking with the forked tongue (Gillingham and Clark 1981). Chemicals on the tongue are then delivered through an opening in the roof of the mouth to the paired VNO that is located between nasal and oral cavity. Snakes use their VNO for prey detection and intraspecific communication (reviewed in Halpern 1987; Halpern and Martínez-Marcos 2003).

1.5 Vomeronasal organ in mammals

The VNO of mammals is a paired organ located at the basal part of the nasal cavity (Døving and Trotier 1998). Two tubular structures are separated by the nasal septum and the vomer bone (Fig. 1.2). The VNO receptor neurons (= sensory epithelium) form a crescent-shape (in sagittal section) with its concave side located laterally. Each crescent encompasses a VNO lumen and a blood vessel. The organ is surrounded by a bony or cartilaginous capsule with a narrow duct to the base of the nasal cavity (nasal duct) in some rodents, most ungulates and lagomorphs (Bertmar 1981; Keverne 1999). In most other mammals with functional VNO (e.g. insectivores, carnivores, most rodents, some ungulates and some primates), the organ is connected with both oral and nasal cavity through the nasopalatine duct (Bertmar 1981). This allows molecules to reach the organ either through the nose or through the mouth, but the duct is vestigial or lost in all aquatic mammals. The VNO lumen is blind-ended – open only to the anterior end – and filled with fluid secreted by vomeronasal glands. The proximal side of the lumen is lined with vomeronasal sensory epithelium that contains bipolar sensory cells similar to sensory cells in the MOE (Retzius 1894). However, these cells have microvilli on their single dendrite in contrast to cilia that are typical for the MOE (Bannister 1968). The axon on the other side of the cell body leads to the AOB. Large blood vessels surrounding the VNO lumen can cause vasodilatation and vasoconstriction leading to an autonomically controlled vascular pumping mechanism for stimulus access (Meredith et al. 1980; Meredith 1994). At
least in hamsters, the pump seems to respond to any novel situation where the animal’s attention is attracted. Because molecules have to be transported to the vomeronasal sensory epithelium in aqueous medium, water solubility might be more important for molecules perceived by the accessory than by the main olfactory system. However, transporter proteins produced by nasal glands are likely to bind odorants or pheromones to assure solubility of the bound complex (e.g., vomeromodulin in rats: Krishna et al. 1995). Several ungulates, cats and some other mammals show a behaviour called “flehmen” that facilitates the inflow of molecules to the VNO (shown in goats by Melese-d’Hospital and Hart 1985). During this behaviour that in ungulates is typically seen when males investigate female secretions, the animal raises its head and curls back the upper lip (review of flehmen in mammals by Estes 1972).

Fig. 1.2: Sagittal section of mouse VNO (middle part of the VNO, after von Mihalkovics 1899, Fig. 52, and Døving & Trotier 1998, Fig. 2)

1.6 Vomeronasal organ in humans

The functionality of the VNO in humans is highly debated (reviewed by Martínez-Marcos 2001; Meredith 2001). In foetuses the VNO starts to develop, but subsequently regresses (e.g., Boehm and Gasser 1993), leading to vestigial or lost VNOs in adults (reviewed by Wysocki 1979). A study in 1985 showed that 39% of adult humans have an opening to the VNO in the nasal cavity (Johnson et al. 1985). Vomeronasal structures were
found in 70% of adults but there were no signs of functionality. Although the VNO is anatomically distinguishable, the organ does not show the typical tubular structure of bone or cartilage. A later study verified that 39% of adult humans have at least one opening (13% two, 26% one) but also showed changes over time in the presence of the vomeronasal pit (Trotier et al. 2000). In ~73% of the tested population at least one pit was clearly defined on some days. However, by using antibodies in vomeronasal epithelial cells no olfactory marker protein could be detected, which is expressed in vomeronasal sensory neurons of other mammals (Trotier et al. 2000). Also no S100 was found, a protein expressed in Schwann cells. Schwann cells would have indicated the existence of vomeronasal nerve bundles connected with the brain. Because of the loss of sensory epithelium and other vomeronasal structures the human VNO is generally considered as non-functional (Bhatnagar and Smith 2010). This is supported by the apparent lack of AOBs in humans (Meisami et al. 1998).

However, an involvement of the MOE rather than the VNO to detect potential pheromones is suggested (also shown in other mammals, see Baxi et al. 2006) as several “pheromonal” effects were described in humans: for example, odourless compounds from the armpits of women can shorten or lengthen the menstrual cycle of recipient women which leads to a synchronisation of the menstrual cycle (McClintock 1971; Stern and McClintock 1998). Furthermore, women have the ability to discriminate and choose male odours (Jacob et al. 2002). The secretions of the areolar glands of lactating women initiated positive head turning and breastfeeding in newborns (Doucet et al. 2009). There is early evidence for real human pheromones to be present in the areolar glands, however, no molecules in humans have been chemically identified yet to which the term “pheromone” can be appropriately applied (Wyatt 2014, pp. 275, 295-296).

1.7 Olfactory and vomeronasal receptors

Olfactory receptors (ORs), which are the main class of receptors in the MOE, were first described in 1991 (Buck and Axel 1991). These receptors are expressed in the cilia of the olfactory sensory neurons and are coupled with GTP-binding proteins (G-proteins). Following ligand binding to the receptor, the α-subunit of the G-protein dissociates and activates adenylate cyclase. The subsequently increasing cAMP (cyclic adenosine monophosphate) concentration opens cyclic nucleotide-gated cation channels which causes a depolarisation of the nerve cell. The action potential is transmitted along the axon to the MOB. It was proposed that the main olfactory system uses combinatorial coding: “one OR recognizes multiple
odorants, and [...] one odorant is recognized by multiple ORs” (Malnic et al. 1999, p. 713). Combinatorial coding could explain why humans, for example, can detect more than 10,000 different odours with only a few hundred OR genes (Niimura 2012). Although ORs are mainly expressed in the MOE, there are some functional ORs in the testis of several mammalian species (Parmentier et al. 1992; Vanderhaeghen et al. 1997; Spehr et al. 2003, not covered in this introduction). The MOE expresses trace amine-associated receptors (TAARs) as a second class of chemosensory receptors (Libers and Buck 2006). TAARs can recognize volatile compounds in mice urine and are therefore suggested to be involved in pheromone detection of the MOE. A third class of receptors, guanylyl-cyclase-D (GC-D), is expressed in a subpopulation of sensory neurons in the rat MOE (Füll et al. 1995).

It was first demonstrated in rabbits that the VNO consists of two populations of nerve cells (Mori et al. 1987). Monoclonal antibodies revealed that one subclass of vomeronasal nerve fibres projects to the glomeruli in the rostrrolateral (~posterior) part of the accessory olfactory bulb, whereas the axons of the other subclass end in the caudomedial (~anterior) part. A similar separation of rostral and caudal accessory olfactory bulb was found in rodents (Schwarting and Crandall 1991; Schwarting et al. 1994; von Campenhausen et al. 1997). The rostral projecting nerve cells have shorter dendrites, are located in the apical region of the vomeronasal sensory epithelium (closer to the VNO lumen) and are coupled with \( G_{\alpha i2} \)-type G-proteins (Halpern et al. 1995; Berghard and Buck 1996). In contrast, the caudal projecting nerve cells have longer dendrites, are located in the basal region of the epithelium (closer to the nasal septum), and are coupled with \( G_{\alpha o} \)-proteins (Fig. 1.3).

Fig. 1.3: G-protein hybridisation signals in the VNO of mice (sagittal section); the white-framed black area indicates the VNO lumen; scale bar 50 pm (after Berghard and Buck 1996, Fig. 1B)
The expression of the two types of G-proteins coincides with the expression of the two main types of vomeronasal receptors found in rodents (Fig. 1.4): Vomeronasal receptors 1 (V1R) are found in apical sensory epithelium (Dulac and Axel 1995), whereas vomeronasal receptors 2 (V2R) show structural differences to V1Rs and are expressed in the basal epithelium of the VNO (Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). The low sequence similarity of V1Rs and V2Rs indicates separate evolutionary origins (Swaney and Keverne 2009). Recently, a third class of VRs – formyl peptide receptors (FPRs), also called FPR-like proteins – was discovered in mice and seems to be rodent-specific (Liberles et al. 2009; Rivière et al. 2009). The FPRs in the VNO are specifically expressed in this organ and are probably used for pathogen sensing (Rivière et al. 2009; Bufe et al. 2012), but other FPRs are expressed in a large variety of cells (Migeotte et al. 2006). V1Rs, V2Rs and FPRs are all part of the large family of seven transmembrane domain G-protein coupled receptors (GPCR) and chemosensory receptor genes, in general, are one of the functional categories that are most frequently found to be under positive selection in a variety of animal taxa (Kosiol et al. 2008).

Fig. 1.4: Vomeronasal receptors V1R and V2R. Some V2Rs are coexpressed and interact with MHC class 1b (M10). Both VRs require the cation channel protein TRPC2 (after Dulac and Torello 2003, Fig. 3b).

V1Rs are encoded by a single exon (Dulac and Axel 1995), whereas V2Rs have complex intron/exon structures and possess a relatively long N-terminal extracellular domain (Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). Subsets of V2R expressing neurons also express different genes of MHC families (Fig. 1.4, Ishii et al. 2003; Loconto et al. 2003). Every neuron in the VNO expresses only one type of VR (Mombaerts 2004). In V1R-expressing cells only one of the two alleles of a gene and no other VR gene is expressed, known as monoallelic expression (Rodriguez et al. 1999; Del Punta et al. 2002b). Only if a transcribed V1R allele is defect, does the cell express a new, functional
gene (Roppolo et al. 2007). Monoallelic expression is also known from ORs in the MOE where it was described first (Chess et al. 1994). V2Rs show coordinated coexpression where each sensory neuron expresses two V2R genes (Ishii and Mombaerts 2011). Here, one gene of three V2R families is coexpressed with a member of the fourth family. Although hundreds of VRs have been classified in mammals, little is known about the corresponding ligands that bind these receptors. In general, V1Rs bind smaller more volatile molecules (Leinders-Zufall et al. 2000; Del Punta et al. 2002a), whereas V2Rs bind larger, often water-soluble peptides like MHC (major histocompatibility complex, Leinders-Zufall et al. 2004), ESP1 (exocrine gland-secreting peptide 1, Kimoto et al. 2005; Haga et al. 2010), and MUPs (major urinary proteins, Chamero et al. 2007). For example, in mice it was shown that V1rb2 expressing cells respond to 2-heptanone (Boschat et al. 2002), a molecule in the urine of pregnant or lactating females that accelerates puberty in other females (Jemiolo et al. 1989). V1rb2 is also an exclusive example as it was reported that it might be expressed in MOE and VNO as 2-heptanone activates both MOB and AOB (Xu et al. 2005). Expression of single V1Rs has actually been shown in some cells of the MOE in mice (Karunadasa et al. 2006) and also goats show the expression of the same V1Rs in both the MOE and VNO (Wakabayashi et al. 2002; Ohara et al. 2009). However, the broader pattern of VR expression is yet unclear.

Another protein besides the VRs that is associated with the VNO and is essential for its functionality is the transient receptor potential channel type C2 (TRPC2, Fig. 1.4). It is only present in the VNO tissue (and in less than 1% of MOE cells, Liman et al. 1999; but no MOE-expression was found by Zhang et al. 2010) and there, its expression is highly restricted to the vomeronasal microvilli (Brann et al. 2002). However, less is known about the expression of TRPC2 outside of rodents.

Genes encoding V1Rs and TRPC2 are expressed in the olfactory organ of jawless fish. Therefore, the genetic components evolved in the common ancestor of all extant vertebrates (Grus and Zhang 2009). In mammals V1R repertoires vary immensely in size across different taxa (Grus et al. 2005; Young et al. 2010). Rodents generally have large V1R repertoires of 89 to 239 estimated intact genes, whereas no intact genes were found in dolphins and bats (Young et al. 2010). Marsupials have almost 100 intact genes and the platypus’ repertoire consists of 283 estimated intact V1R genes, the highest number known so far. The V1R repertoires in primates show the strongest variation within a mammalian order and vary from as high as 214 estimated intact genes to no functional genes at all and they are highlighted in the next section. Most other mammals analysed have between 30 to 100 intact V1Rs.
Unexpectedly, the dog has only 9 intact V1Rs, which has been attributed to a stronger reliance on ORs for pheromonal communication or to VR depletion by domestication (Young et al. 2005). The VR repertoire of mice has been divided into 12 V1R and 4 V2R families (Rodriguez et al. 2002; Yang et al. 2005; Young and Trask 2007). The VR families in mice share >40% amino acid similarity (Rodriguez et al. 2002) and the broad function of several VRs have recently been identified in mice (Isogai et al. 2011): Olfactory cues from bedding of male and female conspecifics as well as a variety of heterospecific animals (from reptiles, birds and mammals) were used to detect sensory activity in a large number of selectively VR-expressing cells. Although bedding contained urine, faeces, saliva, fur and other gland secretions, and no specific ligand molecule could be assigned to single VRs, the study delivered remarkable insights into VR functionality. Specific V1Rs and V2Rs could be connected to the detection of male or female conspecific cues or the detection of heterospecific cues (Isogai et al. 2011). Detected heterospecifics included closely related species, potentially to minimise interspecific courtship or mating, as well as predator species indicating the ability for olfactory predator recognition.

1.8 Vomeronasal organ and vomeronasal receptor repertoire in primates

The accessory olfactory system shows varying degrees of development in different primate clades. Strepsirrhine primates (lemurs and lorisoids) have a highly developed VNO as well as a relatively large accessory olfactory bulb, which is innervated by VNO neurons (Martin 1990, pp. 329-335; Meisami and Bhatnagar 1998). In tarsiers and New World monkeys the VNO and the accessory olfactory bulb are smaller but still functional, whereas in catarrhine primates (Old World monkeys, apes and humans), the VNO begins to develop during ontogeny but subsequently regresses (Martin 1990, pp. 329-335; Smith et al. 2011). No AOB was found in adult catarrhines (Meisami and Bhatnagar 1998).

In primates, little is understood about vomeronasal receptor diversity in different clades. Only the V1R class of vomeronasal receptors has been identified so far. In the human genome only five putative functional V1R genes among several V1R pseudogenes have been revealed (Giorgi et al. 2000; Rodriguez et al. 2000; Rodriguez and Mombaerts 2002). This finding is clearly consistent with the degeneration of the VNO function in humans. The only gene that is known to be expressed at the mRNA level is VNR1, which is interestingly expressed in the MOE in humans and not in the VNO (Rodriguez et al. 2000). Nevertheless, there is no evidence that any VR is functional in humans. Orthologues of human VNR1 have
been identified in apes and New World monkeys (Mundy and Cook 2003), but the first comparison of \( \text{VIR} \) repertoires across primates was done in 2010, which also included information on \( \text{VIR} \) diversity in strepsirrhine primates (Fig. 1.5, Young et al. 2010). Mouse lemurs (\textit{Microcebus} spp.) show the largest repertoire of any primate (214 estimated \( \text{VIR} \)s) and this genus was therefore selected for the studies on the evolution of VRs in this thesis, which had not been analysed before. Other lemur species might also have large repertoires, but the availability of the mouse lemur genome (although yet incomplete with an average coverage of 2x) and the availability of the published \( \text{VIR} \) sequences is crucial for evolutionary analyses. Further advantages of mouse lemurs as a model system to study the evolution of VRs in primates are listed in the next section. \( \text{V2R} \)s, on the other hand, have been reported to be completely degenerated in primates (Young and Trask 2007). However, the study only included macaques, chimpanzees and humans and consequently missed data about strepsirrhines, tarsiers and New World monkeys, i.e., all primates with functional VNO. In contrast to rodents, VNO-specific \( \text{FPR} \)s are suggested to be absent in primates based on the lack of \( \text{FPR} \)-family size expansion (Yang and Shi 2010).

---

**Fig. 1.5:** Phylogenetic reconstruction of selected primates and estimation of functional \( \text{VIR} \) and \( \text{V2R} \) genes in the genomes of primates (based on Young and Trask 2007; and Young et al. 2010, Fig. 2); ? = no information available

Independent genetic evidence for the loss of VNO function during primate evolution came from the finding that the \( \text{TRPC2} \) gene, which is important for the transduction of VR signalling, is non-functional in catarrhines, due to the presence of multiple deletions (Liman and Innan 2003; Zhang and Webb 2003). It has been widely considered that the evolution of trichromatic colour vision in primates decreased the importance of both the main as well as the accessory olfactory systems (Liman and Innan 2003; Zhang and Webb 2003; Gilad et al.
2004; but see Gilad et al. 2007). However, this hypothesis seems to be wrong at least for ORs as the number of functional genes does not differ significantly between New World monkeys, Old World monkeys and hominoids indicating that the loss of ORs pre-dated the acquisition of full trichromatic vision (Matsui et al. 2010). Niimura (2012) considered that the degeneration of OR genes in primates is linked to the well-developed visual system but cannot be explained by trichromatic vision alone. The hypothesis may still be true for vomeronasal functionality since the pseudogenisation of V1Rs and TRPC2 in catarrhines was dated to ~23 million years ago shortly before the split of Old World monkeys and hominoids (Zhang and Webb 2003).

Immediately before the onset of this thesis in early 2010, the first study on VRs in strepsirrhines was published (Young et al. 2010) as mentioned above. The publication estimated the size of V1R repertoires across mammals in a comparative study (including two species of strepsirrhines) rather than going into detail in terms of repertoire organisation, actual expression or evolution. Information on V1Rs was lacking until recently, when a subset of V1Rs was analysed across several strepsirrhine species (Yoder et al. 2014) and the current thesis aimed to fill the gap of knowledge by analysing different aspects of VRs and their evolution in mouse lemurs as the primate representative with the largest known VR repertoire.

1.9 Model system: mouse lemur

Mouse lemurs (Microcebus spp.) are small nocturnal strepsirrhines that live endemically on Madagascar. The genus includes the smallest primates in the world. Mouse lemurs resemble the ancestral primate condition ("small-bodied nocturnal species with a typical tooth-scaper", Martin 1972b; 1995, p. 547) and are an ideal model system to study the evolution of primates.

Mouse lemurs show remarkable interspecific diversity which was only found during the last decade by studies of their molecular phylogeny, morphology and vocalisations (Yoder et al. 2000; Zimmermann et al. 2000; Radespiel et al. 2012; Rasolarison et al. 2013). While only two species were recognized as recently as 16 years ago, there are now 21 species described and still more are likely. Closely related species have parapatric and sympatric distributions, and several biogeographic hypotheses have been proposed to explain the explosive speciation in this group in relation to river systems and Pleistocene climate change.
Mouse lemurs can be divided into two lineages: the grey and the red lineage.

Female mate choice is probably important in their mating system (Radespiel et al. 2002; Craul et al. 2004) and pheromonal communication and scent marking are likely to be under sexual selection (Heymann 2006). The ecological niche of mouse lemurs is the “fine branch niche” in dense tropical forests (Martin 1972a) which in combination with the nocturnal activity makes them rely more on acoustic and olfactory signals than on vision (e.g., Bunkus et al. 2005; Siemers et al. 2007; Zimmermann 2010 and next paragraph).

Mouse lemurs are an ideal model system to study pheromonal communication in primates. The role of odours in their social system is probably one of the best known for any primate (reviewed in Perret 1995). Mouse lemurs possess one of the largest VNOs and accessory bulbs of any primate in relation to body size (Meisami and Bhatnagar 1998). Additionally, functional studies have shown that their accessory olfactory system specifically responds to urine components (Schilling et al. 1990). Mouse lemurs have several scent glands that are used in different scent-marking behaviours using glandular secretions (mouth wiping, anogenital rubbing) and they also perform urine washing (Glatston 1979). During urine washing the animal urinates into one hand and then rubs the urine on the same-side foot. The actual marking is done by running over the substrate, where the procedure may be repeated with the same or alternate side. Olfactory communication seems to play a central role in the regulation of space use, as mouse lemurs show marking behaviour during the dispersal of sleeping groups at the beginning of the night (Braune et al. 2005). They then forage solitarily (Radespiel et al. 1998) and feed on fruits, nectar, various arthropods, homopteran secretions, and tree gum (Radespiel et al. 2006; Joly and Zimmermann 2007; Thorén et al. 2011). Scent-marking is also excessively used in reproductive advertisement together with acoustic cues (Buesching et al. 1998). Females are able to adjust the sex ratio of their offspring in relation to the presence of urinary cues of other females (Perret 1996). A definite role of the VNO in intraspecific behaviour has been described: Males with surgically removed VNO showed reduced sexual behaviour as well as reduced inter-male aggression (Aujard 1997). Mouse lemurs are able to perceive and react to predator odours (Sündermann et al. 2008; Kappel et al. 2011) which might be mediated by the VNO similar to rodents (Isogai et al. 2011). The VNO is anatomically described in detail (Evans and Schilling 1995; Smith and Rossie 2008) which allows the distinction and subsequent removal of the VNO for gene expression analyses.
1.10 Aims & Hypotheses

This thesis consists of four studies that analyse the evolution and present composition of the VR repertoires in mouse lemurs including \textit{V1R}s (study 1) and \textit{V2R}s (study 2). Subsequently, the expression patterns in the olfactory organs (study 3) and intraspecific variation in the genomic repertoire (study 4) are investigated. The results of all studies are discussed in summary in the general discussion of chapter 6.

Study 1 (chapter 2) analysed the selection pressures on the \textit{V1R} repertoire using the genomic data published by Young and colleagues (2010). Variation in the rate of \textit{V1R} evolution is expected: \textit{V1R}s that bind species-specific pheromones which are potentially important in reproductive biology are expected to have evolved rapidly. In contrast, receptors for gender detection are expected to have evolved slowly. The questions that the study aims to answer are listed below:

1A) What monophyletic \textit{V1R} gene clusters can be identified in the grey mouse lemur and have they evolved under positive selection?
1B) How did individual \textit{V1R} loci evolve in multiple mouse lemur species and does this allow investigating evolutionary processes during the diversification of the genus \textit{Microcebus}?
1C) Are differences in the selection pressures of the grey and the red mouse lemur lineage detectable?
1D) What selection pressures are detectable in the \textit{V1R} repertoire of the greater galago (\textit{Otolemur garnettii}) and are there differences in the evolution of \textit{V1R} genes between two clades of strepsirrhine primates?

Study 2 (chapter 3) investigated whether functional \textit{V2R}s are in fact present in primates. Previously, \textit{V2R}s were considered to be completely pseudogenised in primates based on studies in catarrhines (Young and Trask 2007). Information on strepsirrhines, tarsiers and New World monkeys (those primates with functional VNO) was lacking. The study aims to answer the following questions:

2A) Are \textit{V2R}s degenerated across the primate order or did the loss of \textit{V2R} function occurred only later in primate evolution?
2B) Are putative \textit{V2R}s expressed in the VNO of mouse lemurs and therefore likely to be functional?
Study 3 (chapter 4) investigated whether the expression patterns of V1R and V2R loci in the VNO and MOE of the grey mouse lemur differs between sexes, seasons and ages. All samples were also tested for TRPC2 expression and the following questions were addressed:

3A) What expression patterns are present in the VNO?
3B) Are sexual or seasonal differences present?
3C) How does the expression in an infant mouse lemur compare to adult individuals?
3D) Is expression of VRs detectable in the MOE and are those VRs coexpressed with TRPC2?

Study 4 (chapter 5) analysed the genetic variation and substitutions of VRs within a free-living population and in two species of mouse lemur (M. murinus and M. ravelobensis) in order to identify the mode of selection for different loci. The evolutionary and demographic history of both species is estimated and compared with possible established Pleistocene colonisation scenarios. The study had the following aims:

4A) Are differences in the genetic diversity of VRs detectable in both species?
4B) What are the evolutionary histories of VRs in both species?
4C) Is it possible to reconstruct Pleistocene colonisation events in both species based on VRs?

1.11 References

Albone ES. 1984. Mammalian semiochemistry - The investigation of chemical signals between mammals. 1st edition, John Wiley & Sons Ltd., Chichester, UK.


2. Pervasive and ongoing positive selection in the vomeronasal-1 receptor (V1R) repertoire of mouse lemurs

Published in 2012 in Molecular Biology and Evolution 29:3807-3816 (doi: 10.1093/molbev/mss188)
by Philipp Hohenbrink,1,2 Ute Radespiel,1 and Nicholas I. Mundy2

1Institute of Zoology, University of Veterinary Medicine, Hanover, Germany
2Department of Zoology, University of Cambridge, Cambridge, United Kingdom

Associate editor: Willie Swanson

Chemosensory genes are frequently the target of positive selection and are often present in large gene families, but little is known about heterogeneity of selection in these cases and its relation to function. Here, we use the vomeronasal-1 receptor (V1R) repertoire of mouse lemurs (Microcebus spp.) as a model system to study patterns of selection of chemosensory genes at several different levels. Mouse lemurs are small nocturnal strepsirrhine primates and have a large (~200 loci) repertoire of V1R loci that are likely important for intraspecific pheromonal communication and interspecific interactions, for example, recognition of predator cues. We investigated signals and patterns of positive selection among the 105 identified full length V1R loci in the gray mouse lemur and within 7 V1R loci amplified across multiple mouse lemur species. Phylogenetic reconstructions of published sequences revealed at least nine monophyletic clusters of V1Rs in gray mouse lemurs that have diversified since the split between lemurs and lorisoid primates. A large majority of clusters evolved under significant positive selection. Similar results were found in V1Rs of closely related greater galagos. Comparison with function of related V1R clusters in mice suggested a potential relationship between receptor function and strength of selection. Interestingly, most codons identified as being under positive selection are located in the extracellular domains of the receptors and hence likely indicate the position of residues involved in ligand binding. Positive selection was also detected within five V1R loci (=71% of analyzed loci) sequenced from 6 to 10 mouse lemur species, indicating ongoing selection within the genus, which may be related to sexual selection and, potentially, speciation processes. Variation in strength of positive selection on V1Rs showed no simple relationship
to cluster size. The diversity of \(V1R\) loci in mouse lemurs reflects their adaptive evolution and is most likely related to the fundamental relevance of olfactory communication and predator recognition in these primates. Overall, adaptive evolution is the predominant mode of evolution of \(V1R\) loci at all levels, and the substantial heterogeneity in the strength of selection may be related to receptor function.

### 2.1 Introduction

The vomeronasal organ (VNO) is part of the accessory olfactory system of mammals and some other tetrapods (reviewed in Keverne 1999). It is anatomically and physiologically distinct from the main olfactory system and has traditionally been considered to be specialized for pheromone detection, that is, for intraspecific communication. Although certain pheromones are now known to be sensed by the main olfactory epithelium (Restrepo et al. 2004), recent evidence confirmed the substantial function of the VNO in pheromone detection and also showed that the VNO has a major role in the detection of kairomones (interspecific signals that benefit the receiver without benefiting the emitter), such as predator cues (Papes et al. 2010; Isogai et al. 2011).

In mice, the sensory epithelium of the VNO expresses two classes of vomeronasal receptors that are both seven-transmembrane G-protein coupled receptors: vomeronasal-1 receptors (V1Rs, Dulac and Axel 1995) and vomeronasal-2 receptors (V2Rs, Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). The two vomeronasal receptor types are expressed in distinct regions of the VNO (Keverne 1999), and V1R responses to ligands are characterized by high sensitivity and specificity (Leinders-Zufall et al. 2000; Isogai et al. 2011), which is as expected if they have evolved to respond to specific pheromones and kairomones.

Intact \(V1R\) genes have been found in most mammalian orders examined so far, but there is enormous variation in diversity among taxa. More than 200 distinct \(V1R\) loci are estimated in mouse and platypus genomes (Grus et al. 2005; Grus et al. 2007; Young et al. 2010), for example, whereas only eight intact loci have been found in dogs (Grus et al. 2005). In contrast, \(V2R\) genes have only been described so far in rodents and marsupials. Primates show the most extreme variation in the \(V1R\) repertoire of any mammalian order, with more than 200 \(V1R\) loci predicted in mouse lemurs, ~80 loci in galagos, and only five apparently intact loci in humans, which are probably nonfunctional (Giorgi et al. 2000; Young et al. 2010). This variation in primates reflects the importance of the VNO and the accessory
olfactory system and olfactory communication among different groups. This ranges from a well-developed VNO in strepsirrhines (lemurs and lorisoids) to a small but functional VNO in New World monkeys, apes, and humans (Martin 1990), particularly because TRPC2, a cation channel protein required for transduction of VIR signals, is a pseudogene in these groups (Liman and Innan 2003; Zhang and Webb 2003). It has been proposed that pheromones may be under sexual selection in primates (Heymann 2006). Thus, primates are a very interesting group for the study of VNO and vomeronasal receptor evolution.

Chemosensory receptor genes are one of the functional categories that are most frequently found to be under positive selection in a variety of animal taxa (Kosiol et al. 2008). For VIR loci, positive selection has been detected in a few studies. VIR genes in mice and rats showed strong evidence for positive selection in several clusters, with most evidence for selection occurring shortly after the divergence of mouse and rat lineages (Shi et al. 2005). In an early study on VIR genes in primates, evidence for positive selection was found at a small proportion of sites (Mundy and Cook 2003). However, several important issues in relation to evolution and adaptation of VIR repertoires remain unexplored. Particularly in species with a large VIR repertoire such as mouse lemurs or galagos, is positive selection acting on the whole repertoire? It might be expected that some VIR genes or clusters of genes are evolving more rapidly than others, in relation to the rate of change of their ligands. For example, pheromones signaling sex, age, or individual identity (also known as signature mixtures) would be predicted to evolve less rapidly than pheromones that were sexually selected. Similarly, kairomones that are used to identify predators may evolve more rapidly than kairomones for closely related but reproductively incompatible species. This would in turn be predicted to lead to different rates of evolution in the cognate VIR receptors. The presence of large clusters of receptors that have resulted from repeated rounds of gene duplication enables one to examine whether positive selection is stronger in relation to the size of the cluster. Another key issue is whether ongoing selection on VIR genes among closely related species can be detected, because the majority of studies have investigated deeper events in VIR evolution among distantly related species. Most studies have focused on rodents rather than other mammalian orders such as primates. Finally, there is little known about the structure-function relationship of VIRs, including their ligand-binding sites.

Here, we use a strepsirrhine model, the mouse lemur (Microcebus), to investigate selection pressures on VIRs in primates. Mouse lemurs are an ideal model system to study
pheromonal communication in primates, because the role of odors in the social system of these strictly nocturnal primates is well established (reviewed in Perret 1995). Mouse lemurs have several scent glands that are used in different scent-marking behaviors, they perform urine washing (Glatston 1983), and they also show strong behavioral reactions to mammalian predator odors (Sündermann et al. 2008; Kappel et al. 2011). The huge importance of pheromonal communication and predator detection in mouse lemurs is associated with a complex V1R repertoire, and a previous study already identified ~100 intact individual \( V1R \) loci in the 2x gray mouse lemur (\textit{Microcebus murinus}) genome with a prediction of a total repertoire of ~200 loci (Young et al. 2010). Many aspects of mouse lemur ecology, including their small body size, diet, arboreality, and strong reliance on pheromones are thought to be similar to that of ancestral primates. Additionally, mouse lemurs show a high species diversity with at least 19 different species (Mittermeier et al. 2010; Radespiel et al. 2012) that genetically and morphologically split into two clades, the gray and the red mouse lemurs, and pheromones may have played an important role in sexual selection and speciation by reproductive isolation in this group. They are therefore ideal models to investigate interspecies evolution of V1Rs.

We hypothesize that there will be a variation in the rate of V1R evolution in mouse lemurs. At least some \( V1R \) loci are expected to have evolved rapidly in mouse lemurs, such as those that bind species-specific pheromones that are probably important in reproductive biology. Because large numbers of predator-specific V1Rs are present in mice (Isogai et al. 2011) and mouse lemurs suffer from a high predation pressure (Scheumann et al. 2007), we also predict high numbers of kairomone receptors in the \( V1R \) repertoire of mouse lemurs. Such receptors have probably evolved rapidly as well. On the other hand, we expect some \( V1R \)s to evolve slowly, for example, receptors for gender detection which probably bind sulfated steroids similar to mice (He et al. 2008; Nodari et al. 2008). In this study, we test for positive Darwinian selection in the \( V1R \) repertoire of mouse lemurs. First, we identify monophyletic gene clusters in the \( V1R \) repertoire of the gray mouse lemur (\textit{M. murinus}) and test whether all clusters evolved under positive selection and whether there is heterogeneity in strength of selection among \( V1R \) protein domains. Second, we investigate selection on individual \( V1R \) loci from different parts of the repertoire in multiple mouse lemur species to investigate evolutionary processes during the diversification of the genus. Third, we test for differences in selection pressures between the two mouse lemur clades. Finally, we analyze
selection pressures on V1Rs in the greater galago (*Otolemur garnettii*) to compare the evolution of V1R genes in two clades of strepsirrhine primates.

### 2.2 Materials and Methods

**Data collection**

We used 105 of the 107 previously published V1R sequences from the gray mouse lemur (*M. murinus*, supplementary material from Young et al. 2010). For our nomenclature, we took the last three digits of the listing of the sequences published by Young et al. Instead of “micMurV1R6054_TIs: 1562366684...” we are using the more convenient name “Mmur054” (=>Mmur000 to Mmur106; VN1R-Mmur054 on Genbank following human nomenclature). We removed two sequences – *Mmur102*, which has the same amino acid sequence as *Mmur101*, and *Mmur073*, which is not a full length sequence (it is too short to form the first transmembrane helix and it is missing an ATG start codon). The intact sequences were 864–1,008-bp long. We also used published V1R sequences of other Euarchontoglires (nine primate, one tree shrew, two lagomorph, and five rodent species from Young et al. 2010) to identify monophyletic clusters in the gray mouse lemur. Furthermore, we compared the selection signals on V1Rs in mouse lemurs with those on V1Rs of the greater galago (*O. garnettii*). This species is another nocturnal strepsirrhine primate but a member of the sister clade (Lorisiformes) to lemurs. There are 61 intact galago V1R loci published, out of an estimated 78 (Young et al. 2010).

For locus-specific analyses, we sampled up to 10 different mouse lemur species (*M. bongolavensis*, *M. danfossi*, *M. lehilahytsara*, *M. macarthurii*, *M. mittermeieri*, *M. murinus*, *M. myoxinus*, *M. ravelobensis*, and *M. sambiranensis*, App. 2.1 for capture site information). DNA was extracted from ear tissue using a DNeasy Tissue Kit (Qiagen) and a REPLI-gWGA kit (Qiagen). The V1R loci for resequencing were randomly selected from the published V1R repertoire (Young et al. 2010), except for locus *Mmur033*, which was selected because it is known to be expressed in the VNO tissue of mouse lemurs (Talarico M, personal communication). We selected five loci in gene clusters and two nonclustered loci. Corresponding regions were searched in the mouse lemur genome using basic local alignment search tool. Flanking regions were used to design locus-specific external primers (App. 2.2 for primer sequences). All V1R genes were amplified with Bioline Taq (25µl total volume containing 0.75µl MgCl2 [50mM], 0.05µl of each dNTP [25mM], 2.5µl 10x NH4 reaction buffer IV, 1µl of each primer [10µM], 0.1µl Taq DNA polymerase [5U/µl] and 1µl of DNA)
with the following polymerase chain reaction (PCR) conditions: 94°C for 2 min, 40 times
(94°C for 30 s, primer-specific annealing temperature [60–63°C] for 45 s, 72°C for 90s), 72°C
for 5min. PCR products were sequenced on both strands using BigDye Terminator 3.1
(Applied Biosystems) under standard conditions and run on an Applied Biosystems 3500 or
3730xl 96 capillary sequencing machine.

Data analysis

The consensus sequence of single genes was built with the software SeqMan 5.05
(DNASTAR Inc., Madison, WI, USA). These genes and the 105 intact full length V1R
sequences were aligned and analyzed with MEGA 5 (Tamura et al. 2007). Phylogenetic
reconstructions were performed using the neighbor joining method with maximum composite
likelihood model, pairwise deletions, and 500 bootstrap replications in MEGA.

Codon-based site-specific substitution models were used to estimate $d_S/d_S$ ratios ($\omega$)
with codeml in phylogenetic analysis by maximum likelihood (PAML) 4.4 (Yang and
Bielawski 2000; Yang 2007): Model M0 (“null” model – one single average $d_S/d_S$ ratio
among all sites), M1a (“nearly neutral” model – two classes of sites: one with $d_S/d_S < 1$ and
one with a fixed ratio of 1), M2a (“positive selection” model – three classes of sites: one with
$d_S/d_S < 1$, one with a fixed ratio of 1, the third with a ratio > 1) (Wong et al. 2004), M8a
(“modified null” model – eight classes of sites: eight $d_S/d_S$ ratios ranging from 0 to 1, taken
from a discrete approximation of the beta distribution plus one class of sites with $d_S/d_S = 1$)
(Swanson et al. 2003), and M8 ("beta plus omega" model – eight classes of sites from a beta
distribution like in M8a plus additional class of sites with a $d_S/d_S$ ratio ≥ 1). PAML also
estimates the corresponding proportions of sites under these models. Gaps and ambiguous
sites were not removed ("cleandata = 0"). Likelihood ratio tests were used to determine
whether nested models are significantly more likely ($\alpha = 0.05$) than models that do not allow
sites under positive selection. The test statistic (–2 [log likelihood$_1$ – log likelihood$_2$]) was
calculated to compare models M1a versus M2a and M8a versus M8. The Bayes Empirical
Bayes (BEB) analysis in model M2a was used to identify codons under significant positive
selection. The locations of codons under positive selection were estimated by using the
published structure of human VNIR1 (after Saito et al. 1998; Rodriguez et al. 2000) and
placed in three categories (extracellular, transmembrane, and intracellular domains). A
transmembrane hidden Markov model analysis (results not shown, Krogh et al. 2001) used to
estimate the position of transmembrane helices revealed a strong congruence between the
estimated transmembrane domains and the domains detected by the VNIRI alignment. It can therefore be concluded that the location of the transmembrane domains is highly conserved across the VIR repertoire. Because gene conversion can potentially lead to moderate false detection of positive selection among paralogs using PAML (Casola and Hahn 2009), we tested for gene conversion using the program GENECONV 1.81 (Sawyer 1989).

Site models in PAML were used to detect signals of positive selection in the VIR repertoire on several levels: 1) the full repertoire, 2) all monophyletic clusters of VIR loci, and 3) resequenced single loci. We used the following parameters to compare different clusters: overall $\omega$ value of the null model ($\omega$ M0), the $\omega$ value of the third class in the positive selection model ($\omega_3$ M2a), the corresponding proportion of sites in the positive selection model ($p_3$ M2a), and the product of the last two parameters ($\omega_3 \times p_3$ M2a), as an estimate of the overall “strength” of positive selection. For locus-specific analyses, we used sequence data from seven single loci for 6–10 mouse lemur species. Each analysis included one sequence per mouse lemur species except for the gray mouse lemur (M. murinus), the most widespread species, where we used sequences from two individuals, one southern (near Tolagnaro) and one northwestern (in Ankarafantsika National Park), because of their relatively high nucleotide sequence difference. If heterozygous sites were present (1–6 heterozygous positions per sequence), we used DnaSP 5.10 (Librado and Rozas 2009) to reconstruct the different alleles, built a haplotype network with the software Network 4.6 (Bandelt et al. 1999), and conservatively used the allele most closely related to the alleles of the other mouse lemur species for the subsequent analyses. Sequences have been submitted to Genbank under accession numbers JX275694–JX275755. We used the recent mouse lemur phylogeny of Weisrock et al. (2010).

We also used branch models in PAML to test for differences between the gray and the red mouse lemur lineage. The most closely related published locus to each of the seven selected single loci was used as an outgroup sequence with a separate $d_S/d_S$ ratio (“two-branches” model). The likelihood values of the “two-branches” model were tested against a “three-branches” model, where gray (M. murinus from North and South) and red mouse lemur clades (all remaining tested mouse lemur species) had two different $d_N/d_S$ ratios, using a likelihood ratio test. Additionally, the full VIR repertoire of the galago (Otolemur garnettii) was analyzed with site models in PAML as described earlier. Sister gene clusters to mouse lemur clusters were analyzed in separate PAML runs.
We used observed versus expected $\chi^2$ tests to look for significant differences in the distribution of codons under significant positive selection in the receptor protein, based on the proportion of amino acid residues in each category: ~29% extracellular, ~24% intracellular, and ~47% transmembrane. Codons were assigned to the three domains as described earlier, and each expected distribution was calculated individually per gene cluster. Correlations between different parameters of monophyletic clusters or single genes (see earlier) were tested using the Spearman rank correlation. All statistics were performed in Statistica 6.1 (StatSoft Inc., Tulsa, OK, USA).

2.3 Results

Selection on the mouse lemur V1R repertoire

Phylogenetic reconstruction of available V1R data from the gray mouse lemur revealed the organization of the V1R repertoire (Fig. 2.1, see also Young et al. 2010). Most of the published V1R loci formed nine different clusters (I–IX) that remain monophyletic after including published V1R sequences of other Euarchontoglires species (notably from O. garnettii; phylogenetic tree with V1R sequences of primates, tree shrew, lagomorphs, and rodents in supplementary material S3, Supplementary Material online at mbe.oxfordjournals.org; the gray mouse lemur was the only available lemur species). The clusters varied in size from 4 to 29 loci. Ten loci were not clustered and evolved from nine different origins according to the Euarchontoglires tree.

A proportion of gray mouse lemur V1R loci (~16%) showed significant signals for gene conversion (see details in App. 2.5 B), with more than half of putative conversion events occurring in cluster I. The length of the putative converted tracts was generally short (mean: ~15% of the total gene length (~929 bp), range: 67–300 bp). Importantly, PAML analyses run when these conversion tracts were excluded were similar to those before exclusion (App. 2.5 C), and, in particular, in no case was significant positive selection lost. In simulated data with gene conversion, data sets with parameters similar to our data (in terms of $d_S$) showed only a low percentage of false-positive results in site-model tests in PAML (below 5%, Casola and Hahn 2009). We, therefore, conclude that effects of gene conversion on our analysis of selection are minimal and present result without excluding data.
Fig. 2.1: V1R repertoire of gray mouse lemurs (neighbor joining with maximum composite likelihood method, 500 bootstrap replications consensus tree, and bootstrap support within clusters not shown); the roman numbers indicate monophyletic clusters of loci specific to mouse lemurs.
We analyzed selection across the mouse lemur V1R repertoire using a 939 bp alignment (part of the intracellular tail could not be analyzed due to alignment difficulties). The $d_{S}/d_{S}$ ratio of model M0 over all loci was 0.63, and there was significant evidence for a proportion of sites under positive selection over the whole tree in both selection model comparisons (14.6%, Table 2.1). After separate analyses of each monophyletic cluster, seven of nine clusters showed significant evidence for sites under positive selection with 2.2–16.9% of sites selected with $\omega$ from 2.75 to 10.26 (Table 2.1). Only in clusters III and IV was no significant evidence for positive selection found. The number of loci per cluster did not correlate with the overall $\omega$ value of model M0, $\omega_3$ M2a, $p_3$ M2a, or $\omega_3 * p_3$ M2a (Spearman rank correlation, $n = 9$, -0.31 < $R < 0.16$, $P > 0.424$).

Table 2.1: Output of the PAML analysis across the whole V1R repertoire (105 loci) of *Microcebus murinus* and across the separate monophyletic clusters with Likelihood ratio statistics

<table>
<thead>
<tr>
<th>Cluster</th>
<th>No. of Loci</th>
<th>Potential Function</th>
<th>$\omega$ M0</th>
<th>$\omega_3$ M2a</th>
<th>$p_3$ M2a (%)</th>
<th>M1a vs. M2a*</th>
<th>M8a vs. M8*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole repertoire</td>
<td>105</td>
<td>—</td>
<td>0.63</td>
<td>1.89</td>
<td>14.6</td>
<td>106.7***</td>
<td>91.0***</td>
</tr>
<tr>
<td>I</td>
<td>29</td>
<td>—</td>
<td>0.69</td>
<td>3.48</td>
<td>14.0</td>
<td>81.5***</td>
<td>79.4***</td>
</tr>
<tr>
<td>II</td>
<td>7</td>
<td>f h</td>
<td>0.84</td>
<td>3.09</td>
<td>14.7</td>
<td>21.9***</td>
<td>21.8***</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>f h</td>
<td>0.79</td>
<td>1.37</td>
<td>62.3</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>IV</td>
<td>4</td>
<td>f</td>
<td>0.59</td>
<td>2.61</td>
<td>15.5</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>V</td>
<td>14</td>
<td>h</td>
<td>0.81</td>
<td>2.75</td>
<td>16.9</td>
<td>41.4***</td>
<td>41.0***</td>
</tr>
<tr>
<td>VI</td>
<td>9</td>
<td>m f h</td>
<td>0.54</td>
<td>6.51</td>
<td>2.2</td>
<td>12.9**</td>
<td>13.3***</td>
</tr>
<tr>
<td>VII</td>
<td>4</td>
<td>f h</td>
<td>0.99</td>
<td>9.68</td>
<td>8.3</td>
<td>23.9***</td>
<td>23.9***</td>
</tr>
<tr>
<td>VIII</td>
<td>5</td>
<td>f h</td>
<td>0.98</td>
<td>10.26</td>
<td>5.6</td>
<td>18.1***</td>
<td>18.0***</td>
</tr>
<tr>
<td>IX</td>
<td>19</td>
<td>f h</td>
<td>0.64</td>
<td>4.06</td>
<td>11.6</td>
<td>73.7***</td>
<td>73.5***</td>
</tr>
</tbody>
</table>

NOTE — $\omega = d_{S}/d_{S}$, $\omega_3 = d_{S}/d_{S}$ of the third class in M2a, and $p_3$ = proportion of sites in third class of M2a.

a According to mouse orthologs (m, male cues; f, female cues; h, heterospecific cues; Isogai et al. 2011)

b Likelihood ratio test (LRT) (df = 2)

c LRT (df = 1)

** $P < 0.01$

*** $P < 0.001$

**Distribution of sites under positive selection across the V1R protein**

Across the whole V1R repertoire of mouse lemurs, 11 codons were under significant positive selection (BEB $P > 0.95$), and these occurred in just three parts of the protein: the second and third extracellular loops and the seventh transmembrane domain. There was significant over-representation in the extracellular domains (nine codons, Fig. 2.2) compared with transmembrane (two codons) and intracellular domains (no codon) ($\chi^2 = 8.7$, df = 2, $P = 0.013$). The number of positively selected codons in the extracellular domains was significantly higher than expected $\chi^2 = 4.3$, df = 1, $P = 0.038$), whereas the transmembrane domain contained significantly fewer codons under positive selection than expected ($\chi^2 = 4.4$, df = 1, $P = 0.037$).
In the separate analyses on monophyletic clusters, only two clusters had sufficient codons under significant positive selection to test for domain distribution (I and IX), and neither were significant ($P > 0.538$; details about the distribution of codons under significant positive selection in the supplementary material S4, Supplementary Material online at mbe.oxfordjournals.org).

Fig. 2.2: The distribution of positively selected codons in V1R domains. The amino acid sequence of $V1R \text{Mmur101}$ is given as a representative for the whole V1R repertoire; residues under significant positive selection under PAML model M2a are shown in white on black background. The part in brackets was not analyzed across the repertoire due to alignment difficulties. The transmembrane region (TM) is shown in gray; ext, extracellular; int, intracellular.
Selection on individual loci

Using locus-specific primers, we amplified seven individual V1R loci in multiple mouse lemur species, and as there was no indication that more than one locus was amplified per primer pair, we assume we have successfully isolated orthologous loci. Analysis in PAML showed significant evidence for a proportion of sites under positive selection in five of the seven analyzed loci (Mmur033, Mmur048, Mmur060, Mmur066, and Mmur074, Table 2.2). Loci Mmur060 (cluster V), Mmur048 (cluster VI), and Mmur074 (cluster IX) evolved under significant positive selection as did the whole clusters they belong to. However, the selection pressures during the evolution of the single loci were not always consistent with the selection pressures within the corresponding cluster in the gray mouse lemur. Cluster VII showed strong signals for positive selection, but no such evidence was found for locus Mmur045 or Mmur054. The nonclustered single origin loci Mmur033 and Mmur066 evolved under significant positive selection. Individual loci and the corresponding clusters were uncorrelated for $\omega$ M0, $\omega_3$ M2a, $p_3$ M2a, or $\omega_3^* p_3$ M2a (Spearman rank correlation, $n = 5$, $-0.88 < R < -0.05$, $P > 0.053$). Only 1–3 codons under significant positive selection significant sites were found in the other five loci, and because of the small sample size, no statistical analyses were conducted (details about the distribution of codons under significant positive selection in the supplementary material S4, Supplementary Material online at mbe.oxfordjournals.org).

The branch models showed that the $d_S/d_S$ ratios were higher in the gray than in the red mouse lemur clade for all seven loci. At the level of individual loci, one locus was significantly different – Mmur074 ($\chi^2 = 3.9$, df = 1, $P = 0.048$; all other loci with $\chi^2 < 1.2$, $P > 0.290$).

Table 2.2: V1R loci analyzed in multiple mouse lemur species, with length of the analyzed region (full coding sequence length in brackets), number of sampled species (App. 2.1), and output of PAML analyses with Likelihood ratio statistics

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cluster</th>
<th>Analyzed Region (bp)</th>
<th>No. of Species</th>
<th>$\omega$ M0</th>
<th>$\omega_3$ M2a</th>
<th>$p_3$ M2a (%)</th>
<th>M1a vs. M2a*</th>
<th>M1a vs. M2a**</th>
<th>M1a vs. M2a***</th>
<th>M1a vs. M2a***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mmur033</td>
<td>—</td>
<td>942 (942)</td>
<td>10</td>
<td>0.69</td>
<td>7.51</td>
<td>5.2</td>
<td>9.2*</td>
<td>9.2*</td>
<td>9.2*</td>
<td>9.2*</td>
</tr>
<tr>
<td>Mmur045</td>
<td>VII</td>
<td>915 (945)</td>
<td>7</td>
<td>0.65</td>
<td>1.67</td>
<td>40.3</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Mmur048</td>
<td>VI</td>
<td>957 (957)</td>
<td>6</td>
<td>0.58</td>
<td>5.29</td>
<td>13.7</td>
<td>14.6**</td>
<td>14.6**</td>
<td>14.6**</td>
<td>14.6**</td>
</tr>
<tr>
<td>Mmur054</td>
<td>VII</td>
<td>897 (897)</td>
<td>6</td>
<td>0.24</td>
<td>1.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Mmur060</td>
<td>V</td>
<td>906 (906)</td>
<td>9</td>
<td>0.29</td>
<td>5.62</td>
<td>4.2</td>
<td>7.3***</td>
<td>7.3***</td>
<td>7.3***</td>
<td>7.3***</td>
</tr>
<tr>
<td>Mmur066</td>
<td>—</td>
<td>897 (897)</td>
<td>9</td>
<td>0.48</td>
<td>6.92</td>
<td>5.7</td>
<td>13.0**</td>
<td>13.0**</td>
<td>13.0**</td>
<td>13.0**</td>
</tr>
<tr>
<td>Mmur074</td>
<td>IX</td>
<td>918 (918)</td>
<td>8</td>
<td>0.56</td>
<td>4.18</td>
<td>7.9</td>
<td>7.3***</td>
<td>7.3***</td>
<td>7.3***</td>
<td>7.3***</td>
</tr>
</tbody>
</table>

NOTE — $\omega = d_S/d_S$, $\omega_3 = d_S/d_S$ of the third class in M2a, and $p_3 = $ proportion of sites in third class of M2a

* LRT (df = 2)

** LRT (df = 1)

* $P < 0.01$

** $P < 0.001$

*** $P < 0.05$
Comparison with mouse V1R clades and potential function of mouse lemur V1R gene clusters

The phylogenetic reconstruction of V1R evolution in Euarchontoglires (supplementary material S3, Supplementary Material online at mbe.oxfordjournals.org) revealed the relationship between clades of V1Rs in mice (V1Ra–V1Rk) and mouse lemurs. Cluster I of mouse lemurs was distinct, whereas clusters II and III lay within the mouse V1Ref clade. Clusters IV, V, VI, and VII were most closely related to mouse clades V1Rd, V1Rjk, V1Ri, and V1Rh, respectively. Clusters VIII and IX were both most closely related to clade V1Rc. Some mouse V1R clades are specific to one class of odor cue (Isogai et al. 2011). If the general function of V1R clades is conserved, we could use the function of mouse clades to tentatively assign specific functions to closely related mouse lemur clusters (Table 2.1, Isogai et al. 2011). Thus, according to mouse V1R function, mouse lemur V1Rs in cluster IV would detect female cues, cluster V would detect heterospecific cues, and detection of male cues would be coded by cluster VI genes. Interestingly, cluster IV showed no evidence of positive selection, whereas clusters V and VI showed strong evidence of positive selection (Table 2.1). The remaining clusters (instead of cluster I that seems to be mouse lemur specific) would detect female and heterospecific cues.

Comparison with another strepsirrhine

The known functional V1R repertoire of the galago (O. garnettii) consists of 61 different loci, with a further 17 loci predicted (Young et al. 2010), less than half the number found in mouse lemurs. The overall repertoire evolved under significant positive selection (\( \omega_{M0} = 0.54, \omega_{M2a} = 1.74, p_{M2a} = 11.5\%, \) likelihood ratio tests: M1a vs. M2a, df = 2, \( \chi^2 = 49.3, P \leq 0.001; M8a vs. M8, df = 1, \chi^2 = 43.1, P \leq 0.001 \)). In total, six monophyletic clusters were found in the galago, of which five were sister clusters to the mouse lemur clusters I, III, V, VI, and VIII, and hence presumably diverged at the time of the lemur–lorisoid split. Galagos had a further specific cluster closely related to the single locus Mmur066 in mouse lemurs. Cluster sizes were smaller in galagos than in mouse lemurs (maximum of 10 loci per cluster compared with 29 in mouse lemurs). The clusters IV and VII of mouse lemurs had no direct galago sister sequence or cluster, whereas clusters II and IX were closely related to a single galago locus. Comparisons of pairs of closely related sister clusters in galagos and mouse lemurs showed similar results, that is, cluster III and the corresponding cluster in the galago V1R repertoire showed no significant signals of positive selection, whereas the four other mouse lemur clusters and their corresponding galago clusters all evolved under
significant positive selection (detailed results are shown in App. 2.5 A). There was no correlation between sister clusters in the overall ω of M0, ω3 M2a, p3 M2a, or ω3 * p3 M2a (Spearman rank correlation, n = 5, -0.8 ≤ R ≤ 0.6, P > 0.104).

2.4 Discussion

We analyzed patterns of selection on mouse lemur V1R loci at several different levels. We found evidence for pervasive positive selection in the whole V1R repertoire and in multiple monophyletic V1R clusters, as well as evidence for ongoing selection on individual V1R loci within mouse lemurs. In addition, there was some evidence for variable selection pressures on different loci. As now discussed, these results have important consequences for models of V1R evolution. It has to be kept in mind that the analyzed V1R repertoire contains ~50% of the total estimated number. Nevertheless, we have no reason to expect any sampling bias.

Patterns of selection pressures on V1R genes

There was strong evidence for positive selection across the whole mouse lemur V1R repertoire, which echoes the results of studies on V1R selection in rodents (mouse/rat: Shi et al. 2005; mice: Park et al. 2011). We also investigated patterns of selection in monophyletic clades of V1Rs in the gray mouse lemur. Remarkably, seven of nine monophyletic V1R clusters showed evidence of positive selection in site tests, showing that this is the dominant mode of V1R evolution over all clusters. The lack of signals in two clusters (III and IV) could be related to low statistical power relating to the small number of genes per cluster (4 in both cases). However, clusters VII (four loci) and VIII (five loci) showed highly significant signals of selection despite their low sample size, which suggests that clusters III and IV may have rather evolved under purifying selection. The results, therefore, provide some support for our prediction that some V1R loci are evolving more rapidly than others. However, there was no evidence that the strength of selection in a particular cluster was correlated with the size of the cluster, which would be predicted if the rate of gene duplication is related to the strength of selection. Although only sequences from the gray mouse lemur were used for the analyses of the whole V1R repertoire and monophyletic clusters, patterns most likely contain evolutionary information from the whole genus Microcebus and potentially other lemurs.

The comparison with V1R clades of mice revealed the potential function of the gene clusters in mouse lemurs, and several patterns show good concordance with these functions.
Mouse lemur cluster IV is a small cluster with no significant evidence of positive selection, and genes within this cluster should bind female cues. On the other hand, clusters V and VI and two of their individual loci (Mmur060 and Mmur048, respectively) should detect heterospecific or male cues and showed strong evidence for positive selection. Female cues include sulfated steroids (Nodari et al. 2008; Isogai et al. 2011), which are strongly conserved over evolutionary time, whereas some male cues may rapidly evolve by sexual selection (either by male–male competition or female choice), and interspecific predator cues (kairomones) may be evolving in arms races in some circumstances.

We compared the selection on V1Rs between the gray mouse lemur and greater galago, which is a member of the sister group (Lorisiformes) to the lemurs. The two species share strong ecological similarities, being small, nocturnal, and arboreal (Bearder 1987). As in mouse lemurs, the majority of species-specific V1R clusters in the greater galago were under positive selection. The data suggest some conservation of the rates and possibly function of molecular evolution among the two groups, because the selection on sister V1R clusters was broadly similar suggesting some conservation of function among related V1R clusters in different groups, which is consistent with the discussion of mouse lemur V1R clusters above. However, the gray mouse lemur is predicted to have more than twice as many functional V1R loci as the galago, and it is interesting to note this is related to a higher number of V1R loci per cluster in the mouse lemur, rather than an increase in number of clusters.

To investigate the more recent evolution of V1R loci in mouse lemurs, we resequenced a set of single loci from up to 10 divergent mouse lemur species. We found significant signals for positive selection in five out of seven analyzed loci, showing frequent ongoing positive selection in the V1R repertoire in the genus Microcebus. This is one of the first studies to demonstrate positive selection on single V1R loci in a phylogenetic context and shows that selection in V1Rs is not confined to deeper evolutionary events. We found no correlation between the strength of positive selection in a V1R cluster in the gray mouse lemur and the strength of selection on a particular locus in the cluster among Microcebus species but demonstrated heterogeneity in the rate of molecular evolution of V1R loci within clusters. Indeed, the two loci that showed no evidence for positive selection among Microcebus species were members of cluster VII, which has been shown to be under positive selection. So either different V1R loci in the same cluster are evolving differently in mouse lemurs or there are species differences in the evolution of individual loci. However, we found no significant differences between the signals of positive selection in the gray and the red clade in either of
the analyzed cluster VII loci using branch models. An intriguing possibility is that ongoing positive selection in individual loci could be related to reproductive isolation and potentially speciation within the genus *Microcebus*. Isogai et al. (2011) showed that in mice vomeronasal receptors activated by male cues of conspecific and heterospecific mice are often closely related but mostly distinct. As mentioned earlier, *Mmur048* is a member of a cluster (VI) whose sister clade in mice contains V1Rs responding to male cues, so this is a candidate locus involved in sexual selection. One other study has investigated selection on individual V1R loci using a population genetics approach in a few species of mice (Park et al. 2011), with positive selection detected in ~5% loci. However, the different experimental approach, which is assaying more recent selection than in our study, makes quantitative comparisons between the two studies difficult.

Overall, therefore, we found that positive selection has acted in many different V1R lineages in mouse lemurs and galagos, and in different *Microcebus* species. We have identified two loci, *Mmur048* and *Mmur060*, under selection in mouse lemurs that are candidate male pheromone and kairomone receptors, respectively. We also found some evidence of heterogeneity in strength of positive selection among clusters and individual loci. However, the evolutionary background and function of this heterogeneity remains to be elucidated and is an important question for future research.

**Selection on protein level**

We investigated the distribution of positively selected sites within the V1R protein. For the whole V1R repertoire of mouse lemurs, amino acid residues with significant evidence for an $\omega$ above 1 were significantly biased toward the extracellular domain and specifically toward the third and fourth extracellular domain (= second and third extracellular loop). Little is known about the ligand binding domains of V1Rs, and these results strongly suggest that the second and third extracellular loops are the most important for affecting ligand binding. A study on mice and rats found only single sites under positive selection in different subsets of V1R sequences (Shi et al. 2005). One subset, however, showed a handful of significant sites located within the third extracellular loop (positions 262, 263, and 266) or close to transmembrane regions (259 and 270) after aligning with our data.

In conclusion, V1Rs in mouse lemurs evolved under strong but variable positive selection reflecting the importance of pheromone communication and kairomone detection in these primates. As nonclustered loci were also targeted, it may be possible to find positive
selection in other primate species that may not have such a diversified V1R repertoire. Future studies on V1Rs should therefore consider species with both larger and smaller repertoires. Further research is needed to understand the role of vomeronasal receptors in primates.

The authors thank S. Montgomery and H. Breitrück for their help in the laboratory. They also thank the Institute of Zoology of the University of Veterinary Medicine Hanover (Prof. E. Zimmermann) for logistical support and reviewers for useful comments on the manuscript. This work was funded by the Volkswagen Foundation (grant number I/84 798).

2.5 References


App. 2.1: Capture sites and data for all mouse lemur species. GPS coordinates are given. For each V1R locus, an X indicates if data were obtained.

<table>
<thead>
<tr>
<th>Species</th>
<th>Capture site</th>
<th>GPS</th>
<th>Mmur033</th>
<th>Mmur045</th>
<th>Mmur048</th>
<th>Mmur054</th>
<th>Mmur060</th>
<th>Mmur066</th>
<th>Mmur074</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. bongolavensis</td>
<td>Maroakata</td>
<td>16°04'56.60&quot;S, 47°18'04.80&quot;E</td>
<td>X</td>
<td>X</td>
<td>X***</td>
<td>*</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. danfossi</td>
<td>Ambarjeby</td>
<td>14°56'29.20&quot;S, 47°42'44.50&quot;E</td>
<td>X</td>
<td>**</td>
<td>X</td>
<td>*</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. lehilahytsara</td>
<td>Mantadia</td>
<td>18°49'60.00&quot;S, 48°25'35.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. macarthurii</td>
<td>Anjahely Makira</td>
<td>15°24'22.50&quot;S, 49°29'54.30&quot;E</td>
<td>X</td>
<td>**</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. mamiatra</td>
<td>Manehoko</td>
<td>13°25'49.00&quot;S, 48°47'51.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. mittermeieri</td>
<td>Anjahely Makira</td>
<td>15°24'45.00&quot;S, 49°29'37.50&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. murinus (North)</td>
<td>JBA Ankarafantsika</td>
<td>16°19'00.00&quot;S, 46°48'00.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. murinus (South)</td>
<td>Mandena</td>
<td>24°47'00.00&quot;S, 47°08'60.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. myoxinus</td>
<td>Madirovalo</td>
<td>16°22'46.60&quot;S, 46°29'02.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. ravelobensis</td>
<td>JBA Ankarafantsika</td>
<td>16°19'00.00&quot;S, 46°48'00.00&quot;E</td>
<td>X</td>
<td>X</td>
<td>X***</td>
<td>*</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M. sambiranensis</td>
<td>Mahilaka</td>
<td>14°17'12.00&quot;S, 48°13'38.00&quot;E</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = primers for Mmur054 did not seem to work for the M. bongolavensis/M. danfossi/M. ravelobensis subclade
** = Mmur045 locus in M. macarthurii and M. danfossi contained one or two stop codons respectively inside the coding sequence, indicating possible recent pseudogenization
*** = coding sequence for Mmur048 in M. bongolavensis and M. ravelobensis were identical; sequence was only included once

App. 2.2: List of primers to amplify specific V1R genes in mouse lemurs (continues on next page)

<table>
<thead>
<tr>
<th>Locus name</th>
<th>Genebank ID of locus</th>
<th>Primer name</th>
<th>Primer sequence (5'-3')</th>
<th>Anneal. temp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mmur033</td>
<td>ABDC01467686</td>
<td>Mmur033 fw</td>
<td>AGTGACTGGGAGACTGGCAGA</td>
<td>60.9°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur033 rv2</td>
<td>TGAATAACCAACTGAACATTGTG</td>
<td>57.6°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur033 INT-fw</td>
<td>AGCAGTGCCAACACATCCAT</td>
<td>60.4°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur033 INT-rv</td>
<td>TCGTGATCACCTTGGGAAGACA</td>
<td>60.2°C</td>
</tr>
<tr>
<td>Mmur045</td>
<td>ABDC01423622</td>
<td>Mmur045 fw2</td>
<td>TTTAGAGATCTAAAGGAGCAAT</td>
<td>57.1°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur045 rv</td>
<td>GCAGGTCAACATGTTTGTG</td>
<td>60.0°C</td>
</tr>
<tr>
<td>Locus name</td>
<td>Genebank ID of locus</td>
<td>Primer name</td>
<td>Primer sequence (5’-3’)</td>
<td>Anneal. temp.</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>-------------</td>
<td>--------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Mmur048</td>
<td>ABDC01319044</td>
<td>Mmur048 fw</td>
<td>TCCCAGTTTTGGCTATGATG</td>
<td>59.9°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur048 rv</td>
<td>ACTGCTAATTITAGTGAGGAAATGT</td>
<td>59.4°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur048 INT-fw</td>
<td>TGCCTCTACCTCAAGTTTATGTA</td>
<td>60.1°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur048 INT-rv</td>
<td>GTAGGAGACAGTGGGATACAGA</td>
<td>59.5°C</td>
</tr>
<tr>
<td>Mmur054</td>
<td>ABDC01381893</td>
<td>Mmur054 fw</td>
<td>CATCCTCAATTACTAGTGGCTTTCC</td>
<td>57.8°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur054 rv</td>
<td>GGGAACATGAAACACAGCGTT</td>
<td>59.7°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur054 INT-fw</td>
<td>GTGCTATATGATGATTTCTTTC</td>
<td>59.8°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur054 INT-rv</td>
<td>CATAGGAGACCTGGGATACAA</td>
<td>59.7°C</td>
</tr>
<tr>
<td>Mmur060</td>
<td>ABDC01417980</td>
<td>Mmur060 fw</td>
<td>GACTGAGATACCTAACAAACATC</td>
<td>58.5°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur060 rv</td>
<td>AGCCATTGTGCTCTTTTGA</td>
<td>59.4°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur060 INT-fw</td>
<td>AGCCAGATCTCACAGACAG</td>
<td>59.9°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur060 INT-rv</td>
<td>GATGGAATAGCTGGGACA</td>
<td>60.0°C</td>
</tr>
<tr>
<td>Mmur066</td>
<td>NT_165853.2/Gl:297498982</td>
<td>Mmur066 fw</td>
<td>CGTCGAGCTTTTCTTGGAG</td>
<td>58.0°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur066 rv</td>
<td>TGAGGACACTAGTAATTTGA</td>
<td>59.8°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur066 INT-fw</td>
<td>CCAGTGTCTACATGGTACTTTCTTCTT</td>
<td>60.0°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur066 INT-rv1</td>
<td>CAAGTGTGCAAGAGGAAGG</td>
<td>60.0°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur066 INT-rv2***</td>
<td>CTGAGGCGGGATAGACAGA</td>
<td>59.9°C</td>
</tr>
<tr>
<td>Mmur074</td>
<td>ABDC01430975</td>
<td>Mmur074 fw</td>
<td>TCTTGATGCCAAAACAAAATA</td>
<td>59.5°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 rv</td>
<td>CTAGGAACACCCATTTTTCAG</td>
<td>60.0°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 INT-fw</td>
<td>CCAGACCAAGAAGCTGAGTT</td>
<td>59.5°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 INT-rv</td>
<td>TGGTCTGAGTCAATTAGA</td>
<td>58.9°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 b077-A**</td>
<td>CTTCTTCTCTCTCTTCCA</td>
<td>59.7°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 b098-G***</td>
<td>CATCCTCAACCTTTTCTGGA</td>
<td>59.7°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mmur074 b793-T** ***</td>
<td>GACCAACAAGYGCCTGGAA</td>
<td>63.4 / 58.1°C</td>
</tr>
</tbody>
</table>

fw = forward
rv = reverse
INT = internal (sequencing) primer
* = internal primers of Mmur054 were also used for Mmur045
** = needed to amplify locus in M. murinus (use fw + b793-T and b077-A + rv in two separated PCRs)
*** = needed to amplify locus in M. lehilahytsara and M. mittermeieri (use fw + b793-T and b098-G + rv in two separated PCRs)
**** = Mmur066 INT-rv2 used in sequencing instead of Mmur066 rv
App. 2.5 A: Output of the PAML analysis across the whole V1R repertoire (61 loci) of *Otolemur garnetti* and across the separate monophyletic clusters with likelihood ratio tests (= LRT); number loci per cluster is shown, nomenclature of cluster after corresponding sister cluster in *Microcebus* (one cluster in *O. garnetti* is represented by a single locus *Mmur066* in *Microcebus*); $\omega = d_N/d_S$, $\omega_3 = d_N/d_S$ of the third class in M2a, $p_3 = \text{proportion of sites in third class of M2a}$

<table>
<thead>
<tr>
<th>Cluster</th>
<th># of loci</th>
<th>$\omega$ M0</th>
<th>$\omega_3$ M2a</th>
<th>$p_3$ M2a</th>
<th>M1a vs. M2a$^a$</th>
<th>M8a vs. M8$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole repertoire</td>
<td></td>
<td>0.54</td>
<td>1.74</td>
<td>11.5%</td>
<td>49.3***</td>
<td>43.1***</td>
</tr>
<tr>
<td>I</td>
<td>6</td>
<td>0.55</td>
<td>7.63</td>
<td>0.4%</td>
<td>20.2***</td>
<td>20.2***</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>0.79</td>
<td>1.99</td>
<td>13.9%</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>V</td>
<td>9</td>
<td>0.68</td>
<td>3.53</td>
<td>8.1%</td>
<td>32.2***</td>
<td>29.5***</td>
</tr>
<tr>
<td>VI</td>
<td>3</td>
<td>1.13</td>
<td>7.43</td>
<td>9.5%</td>
<td>10.2**</td>
<td>10.1**</td>
</tr>
<tr>
<td>VII</td>
<td>7</td>
<td>0.55</td>
<td>4.63</td>
<td>2.6%</td>
<td>8.3*</td>
<td>3.3</td>
</tr>
<tr>
<td><em>Mmur066</em>c*</td>
<td>10</td>
<td>0.62</td>
<td>2.43</td>
<td>16.9%</td>
<td>26.8***</td>
<td>28.5***</td>
</tr>
</tbody>
</table>

$^a$ = LRT (df = 2), $^b$ = LRT (df = 1), $^c$ = this information has been corrected in this thesis and deviates from the published manuscript, where the analysed cluster was wrongfully named “IX”, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$

App. 2.5 B: Estimates gene conversion events using the program GENECONV; corresponding clusters, pair of genes with significant signs of gene conversion, Bonferroni-correlated (BC) p-values and the length of conversion tracts is shown

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Pair showing conversion</th>
<th>BC p-value</th>
<th>Length [bp]</th>
</tr>
</thead>
<tbody>
<tr>
<td>III</td>
<td><em>Mmur037 - Mmur051</em></td>
<td>0.000</td>
<td>143</td>
</tr>
<tr>
<td>V</td>
<td><em>Mmur008 - Mmur063</em></td>
<td>0.000</td>
<td>110</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur091 - Mmur103</em></td>
<td>0.000</td>
<td>177</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur088 - Mmur099</em></td>
<td>0.001</td>
<td>201</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur082 - Mmur103</em></td>
<td>0.001</td>
<td>193</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur050 - Mmur088</em></td>
<td>0.002</td>
<td>179</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur103 - Mmur105</em></td>
<td>0.003</td>
<td>174</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur084 - Mmur088</em></td>
<td>0.006</td>
<td>176</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur088 - Mmur104</em></td>
<td>0.007</td>
<td>184</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur002 - Mmur104</em></td>
<td>0.007</td>
<td>217</td>
</tr>
<tr>
<td>VI</td>
<td><em>Mmur011 - Mmur023</em></td>
<td>0.028</td>
<td>111</td>
</tr>
<tr>
<td>III</td>
<td><em>Mmur020 - Mmur067</em></td>
<td>0.037</td>
<td>75</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur084 - Mmur105</em></td>
<td>0.037</td>
<td>166</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur088 - Mmur093</em></td>
<td>0.041</td>
<td>171</td>
</tr>
<tr>
<td>IX</td>
<td><em>Mmur029 - Mmur079</em></td>
<td>0.045</td>
<td>300</td>
</tr>
<tr>
<td>I</td>
<td><em>Mmur088 - Mmur103</em></td>
<td>0.046</td>
<td>160</td>
</tr>
<tr>
<td>V</td>
<td><em>Mmur019 - Mmur059</em></td>
<td>0.050</td>
<td>67</td>
</tr>
</tbody>
</table>
App. 2.5 C: Output of the PAML analyses across the separate monophyletic clusters before and after exclusion of conversion tracts from the whole alignment, with likelihood ratio tests (= LRT); results without exclusion are also shown in Table 2.1 but are shown again for better comparison; $\omega = d_N/d_S$, $\omega_3 = d_N/d_S$ of the third class in M2a, $p_3 =$ proportion of sites in third class of M2a.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Results without exclusion of data</th>
<th>Results with exclusion of conversion tracts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\omega$ M0</td>
<td>$\omega_3$ M2a</td>
</tr>
<tr>
<td>I</td>
<td>0.69</td>
<td>3.48</td>
</tr>
<tr>
<td>II</td>
<td>0.84</td>
<td>3.09</td>
</tr>
<tr>
<td>III</td>
<td>0.79</td>
<td>1.37</td>
</tr>
<tr>
<td>IV</td>
<td>0.59</td>
<td>2.61</td>
</tr>
<tr>
<td>V</td>
<td>0.81</td>
<td>2.75</td>
</tr>
<tr>
<td>VI</td>
<td>0.54</td>
<td>6.51</td>
</tr>
<tr>
<td>VII</td>
<td>0.99</td>
<td>9.68</td>
</tr>
<tr>
<td>VIII</td>
<td>0.98</td>
<td>10.26</td>
</tr>
<tr>
<td>IX</td>
<td>0.64</td>
<td>4.06</td>
</tr>
</tbody>
</table>

$^a$ = LRT (df = 2), * = $p$≤0.05, ** = $p$≤0.01, *** = $p$≤0.001
3. First evidence for functional vomeronasal 2 receptor genes in primates

Published in 2013 in Biology Letters 9:20121006 (doi:10.1098/rsbl.2012.1006) by Philipp Hohenbrink1,2, Nicholas I. Mundy2, Elke Zimmermann1 and Ute Radespiel1

1Institute of Zoology, University of Veterinary Medicine, Buenteweg 17, 30559 Hanover, Germany
2Department of Zoology, University of Cambridge, Cambridge, UK

Two classes of vomeronasal receptor genes, V1R and V2R, occur in vertebrates. Whereas, V1R loci are found in a wide variety of mammals, including primates, intact V2R genes have thus far only been described in rodents and marsupials. In primates, the V2R repertoire has been considered degenerate. Here, we identify for the first time two intact V2R loci in a strepsirrhine primate, the grey mouse lemur (Microcebus murinus), and demonstrate their expression in the vomeronasal organ. Putatively functional orthologues are present in two other strepsirrhines, whereas, both loci are pseudogenes in a range of anthropoid species. The functional significance of the loci is unknown, but positive selection on one of them is consistent with an adaptive role in pheromone detection. Finally, conservation of V2R loci in strepsirrhines is notable, given their high diversity and role in MUP and MHC detection in rodents.

3.1 Introduction

The vomeronasal organ (VNO) in mammals is an olfactory organ specialized for the perception of pheromones and predator cues (kairomones) (Keverne 1999; Papes et al. 2010). Whereas, rodents possess large VNOs and huge repertoires of both classes of vomeronasal receptors (V1Rs and V2Rs), the size and functionality of the VNO vary immensely in primates. It is well developed in strepsirrhines (lemurs and lorisoids), smaller but still functional in tarsiers and most New World monkeys, and completely non-functional in Old World monkeys, apes and humans (Martin 1990; Smith et al. 2011). Correspondingly, the gene repertoire of V1Rs is huge in strepsirrhines (78 and 214 estimated intact genes in galago and mouse lemur, respectively; Young et al. 2010), but only contains a few putative functional genes in anthropoid primates (e.g. five in humans although not expressed in the
VNO) (Rodriguez and Mombaerts 2002; Young et al. 2010). In contrast, to date, functional V2Rs have only been identified in rodents and marsupials, and the V2R repertoire in primates was reported to be degenerate (Young and Trask 2007). However, the multiple exonic structure of V2Rs in comparison with single exon V1Rs has been considered a hindrance to their identification in databases, and previous studies have only reported catarrhine primate genomes (macaque, chimpanzee and human; Young and Trask 2007). Here, for the first time, we search for V2R genes in available primate genomes with a primary focus on strepsirrhine primates to determine whether V2Rs are indeed degenerate across the primate order or whether loss of V2R function occurred only later in primate evolution. We also use expression of putative V2Rs in the grey mouse lemur to confirm probable functionality, and investigate patterns of natural selection on putative functional loci.

3.2 Material and methods

Data collection

We used BLAST (Basic Local Alignment Search Tool) to search for V2R loci in the three available strepsirrhine genomes (Otolemur garnettii, Genbank AAQR 0000 0000.3; Daubentonia madagascariensis, AGTM 0000 0000.1; Microcebus murinus, ABDC 0000 0000.1). All exons (usually six) of the 122 intact mouse V2R genes derived from the UCSC Genome Browser (Kent et al. 2002) were used in BLAST searches. For potentially functional V2R genes (open reading frame, no stop codon), we used exon 6 to search for orthologues in the 12 further primate species with available genome data (Tarsius syrichta, Callithrix jacchus, Saimiri boliviensis, Macaca fascicularis, Macaca mulatta, Papio anubis, Nomascus leucogenys, Pongo abelii, Gorilla gorilla, Pan paniscus, Pan troglodytes and Homo sapiens). Exon 6 encodes the transmembrane region, which is highly conserved compared with the N-terminal of the receptor (Yang et al. 2005), and has been the main focus in other studies (Zhang et al. 2010). Moreover, exon 5 was identified whenever exon 6 was available, either because it was found on the same shotgun read or in the neighbouring read (Genbank accession numbers in App. 3.3). Sequences were aligned and analysed in MEGA v. 5 (Tamura et al. 2007). Phylogenies were reconstructed using neighbour joining trees with maximum composite likelihood.
Confirmation of transcription

We extracted total RNA from the VNO of a male grey mouse lemur (M. murinus) to confirm the expression of V2R genes in a strepsirrhine primate. The animal was euthanized for veterinary reasons; the VNO was removed immediately after death and stored in RNA later (Qiagen). RNA was extracted with the RNeasy Micro Kit (Qiagen) and transcribed into cDNA using a Quantitect Reverse Transcription Kit (Qiagen) and N6 primer, according to manufacturer’s instructions. RT-PCR and sequencing of TRPC2, which encodes a cation channel protein expressed almost exclusively in the VNO, were used to confirm successful cDNA synthesis from VNO tissue. Several pairs of internal primers on different exons and two external primers (close to the 5’- and 3’-end of exon 1 and exon 6, respectively), were designed to amplify the complete V2Rs (total length of more than 2400 bp) in several amplicons with MyTaq DNA polymerase (Bioline) using sequence information from the three strepsirrhine data (for PCR details, see App. 3.1). PCR products were sequenced on both the strands using BigDye Terminator v. 3.1 (Applied Biosystems) under standard conditions and run on an Applied Biosystems 3500 capillary sequencing machine. Consensus sequences of single genes were built with SEQMAN v. 5.05 (DNASTAR Inc., Madison, WI, USA). Sequences were aligned and analysed using MEGA v. 5.

Mode of selection on V2R genes

We used codon-based site-specific substitution models (codeml in PAML v. 4.4) (Yang 2007) to estimate \( d_N/d_S \) ratios (non-synonymous versus synonymous substitution rates) and to test for positive selection in V2R genes in strepsirrhines. The likelihood of model M1a that does not allow positive selection was compared with that of model M2a that allows positive selection (\( d_N/d_S > 1 \)) using a likelihood ratio test. We also used branch models to estimate different \( d_N/d_S \) ratios in separate primate branches (for details on all PAML analyses, see studies of Hohenbrink et al. 2012).

3.3 Results and discussion

V2Rs in strepsirrhines and tarsiers

BLAST searches revealed two potentially functional V2R genes (greater than 80% sequence, intact open reading frame) in the three strepsirrhine genomes, which we name VN2R1 and VN2R2 (Table 3.1, more details in App. 3.2). VN2R1 was also found to be potentially intact in the Philippine tarsier, T. syrichta, whereas, VN2R2 was a pseudogene in
this species with a frameshift and stop codon in exon 2. Data were missing from some exons, notably exon 4 of VN2R2 which is very short (22 bp).

The expression of both genes in the VNO of the mouse lemur was confirmed by RT-PCR, and full-length cDNAs with intact reading frames were obtained by sequencing (Genbank accession nos KC208006 and KC208007). Thus, at least two intact V2R genes were conserved in strepsirrhine lineages, and these are most likely functional in the VNO. This may be an underestimate of the total number of intact V2Rs (for example, gaps in the genome of the grey mouse lemurs mean that 105 intact V1Rs were found out of an estimated 214; Young et al. 2010), but the total is still very small compared with the large V2R repertoire in mice (greater than 120 loci), showing much lower importance of V2Rs in strepsirrhines. It is also small when compared with the V1R repertoire in strepsirrhines, which have become the dominant class of vomeronasal receptors.

Phylogenetic reconstructions show that VN2R1 groups with mouse V2R family C, whereas, VN2R2 belongs to V2R family D (Fig. 3.1). No strepsirrhine members of mouse families A or B was found, which is interesting since family A is by far the most diverse in mice containing over 80 per cent of all V2Rs. In mice, one V2R gene of families A, B or D is generally coexpressed with one V2R gene of family C in non-random combinations (Silvotti et al. 2007), suggesting the possibility that the two strepsirrhine V2Rs are themselves coexpressed.

Selection analyses revealed that a proportion of codons of VN2R2 significantly evolved under positive selection in strepsirrhines (model M2a $\omega_3 = 22.82$, $p_3 = 1.4\%$; likelihood ratio statistics, LRS = 6.9, df = 2, $p = 0.031$), whereas, VN2R1 did not show signs of positive selection (LRS = 0.0, $p = 1.0$; detailed results can be seen in App. 3.4).

Table 3.1: Sequences for two V2R genes with number of available base pairs from available genome sequences (percentage of total length in round brackets); mouse sequences are Vmn2r1 and Vmn2r56, respectively; Microcebus sequence information was obtained directly by Sanger sequencing; second line shows obtained full-length exons (incomplete exons in square brackets); superscript ‘ps’ denotes exon 2 sequence pseudogenized.

<table>
<thead>
<tr>
<th>gene</th>
<th>Mus</th>
<th>Microcebus</th>
<th>Daubentonia</th>
<th>Otolemur</th>
<th>Tarsius</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V2R1</td>
<td>2739 bp (100)</td>
<td>2739 bp (100)</td>
<td>2319 bp (85)</td>
<td>2739 bp (100)</td>
<td>2281 bp (83)</td>
</tr>
<tr>
<td></td>
<td>exon 123456</td>
<td>exon 123456</td>
<td>exon 123456</td>
<td>exon 123456</td>
<td>exon 123456</td>
</tr>
<tr>
<td>V2R2</td>
<td>2340 bp (100)</td>
<td>2418 bp (100)</td>
<td>1927 bp (82)</td>
<td>2288 bp (98)</td>
<td>1145 bp (52)</td>
</tr>
</tbody>
</table>

 superscript ‘ps’ denotes exon 2 sequence pseudogenized.
Degeneration of VN2R1 and VN2R2 in anthropoid primates

We found homologues of exons 5 and 6 of VN2R1 in 11 anthropoid primate species spanning New World monkeys, Old World monkeys, apes and humans, but all of these were pseudogenes, with at least one exon containing reading frame shifting deletions and/or nonsense mutations. The phylogenetic reconstruction of these sequences recapitulates the primate species phylogeny (Fig. 3.2, Perelman et al. 2011).

The estimated $d_s/d_\sigma$ ratio in branch tests was close to 1 in the anthropoid branch ($d_s/d_\sigma = 0.94$) as expected for pseudogene evolution, while $d_s/d_\sigma$ was lower in the mouse (0.28), the strepsirrhine branch (0.53) and the tarsier (0.35), indicating purifying selection as the main mode of selection that is conserving the function of these V2R genes. For VN2R2, we found further orthologous sequences only in orangutans, both chimpanzee species and humans (all pseudogenes; results not shown). Because only one of the two V2R genes was intact in tarsiers, it appears that pseudogenization of V2R genes began before the split of tarsiiformes.
and anthropoids and continued in the anthropoid stem lineage. Interestingly, both genes are pseudogenized in New World monkeys, although, this clade still possesses a functional VNO in most genera (including *Callithrix* and *Saimiri*, Martin 1990; Smith et al. 2011). We suggest dissimilar importance of V1Rs and V2Rs through primate evolution with V2Rs having started to pseudogenize earlier than V1Rs.

Fig. 3.2: Phylogenetic reconstruction of the combined exons 5 and 6 of VN2R1 from 15 primate species and the mouse (Vmn2r1) as outgroup sequence (neighbour joining tree with maximum composite likelihood method and 1000 bootstrap replications; bootstrap values greater than 50 are shown).

**Functional significance of V2Rs in strepsirrhines**

The conservation of two V2Rs, including positive selection in one of them, strongly suggests important functions in strepsirrhines. In rodents, V2Rs bind mostly non-volatile molecules and have numerous functions. Intriguingly, these include detection of several of the most important classes of pheromones in rodents, including MUPs (major urinary proteins, Chamero et al. 2007; Papes et al. 2010), ESP1 (exocrine gland-secretory peptide 1, Kimoto et al. 2005; Haga et al. 2010) and MHC (major histocompatibility complex) class I peptides (Leinders-Zufall et al. 2004). In addition to intraspecific communication, mice can also detect MUPs from predators using V2Rs (Papes et al. 2010; Isogai et al. 2011). It is known that mouse lemurs react to olfactory cues from predators (Sündermann et al. 2008; Kappel et al. 2011) and females exercise post-copulatory mate choice based on MHC constitution (Schwensow et al. 2008). Given their conservation over millions of years of evolution and the
importance of pheromonal communication in strepsirrhines (DelBarco-Trillo et al. 2011), we can speculate that the intact V2Rs in strepsirrhines likely have an important role in pheromone detection. In conclusion, detection of V2Rs in vertebrate genomes appears to be more straightforward than commonly assumed. Expression data from VNO tissue are important corroborative evidence for V2R (or V1R) function.

We thank the Institute of Animal Breeding at the University of Veterinary Medicine, Hanover for their technical help. We also thank three anonymous referees for their comments and the sequencing centres for generating and releasing their data. This work was funded by the Volkswagen Foundation (grant no. I/84 798).

3.4 References


3.5 Appendix

App. 3.1: PCR mastermix, PCR conditions and primer (primer pairs are shown together); * = external primer sequence, fw = forward, rv = reverse; continues on next page

<table>
<thead>
<tr>
<th>PCR mastermix (Bioline, 25 μl total volume)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 μl MyTaq Reaction Buffer</td>
</tr>
<tr>
<td>1 μl of each primer [10 μM]</td>
</tr>
<tr>
<td>0.1 μl MyTaq DNA polymerase [5 U/μl]</td>
</tr>
<tr>
<td>1 μl of DNA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>94°C for 2 min</td>
</tr>
<tr>
<td>40 times (90°C for 20 s, 60°C for 30 s, 72°C for 60 s)</td>
</tr>
<tr>
<td>72°C for 5 min</td>
</tr>
</tbody>
</table>
Primers (all 5'-3')

VN2R1 5pr fw* TCATTGCCCAAACTCAC
VN2R1 Ex1 rv GATGGTCCTGGAGTGAATAGG
VN2R1 Ex3 fw CGGTTGCTGGCAAGTT
VN2R1 Ex3 rv2 GCAAATCCAATACTTCACCA
VN2R1 Ex3 rv GCAAATCCAATACTTCCACCA
VN2R1 3pr rv* ACTGCAGTTTTGAAGGACACAT

VN2R1 Ex1 rv GATGGTCCTGGAGTGAATAGG
VN2R1 Ex3 fw CGGTTGCTGGCAAGTT
VN2R1 Ex3 rv2 GCAAATCCAATACTTCACCA
VN2R1 Ex3 rv GCAAATCCAATACTTCCACCA
VN2R1 Ex4 rv2 CCACCTTCACAAAGGCAAT
VN2R1 Ex4 fw CAACCCATGATGGAAGAAGAA
VN2R1 Ex6 rv GCAAAGAGAAAAGCCCAGTG
VN2R1 Ex6 fw TCATTGGCAAGCCATCC
VN2R1 Ex6 rv GCAAAGAGAAAAGCCCAGTG
VN2R1 3pr rv* ACTGCAGTTTTGAAGGACACAT

App. 3.2 A: Exons 1 to 6 of the VN2R1 gene in four primate species and the mouse (Vmn2r1) as reference; numbers show base pairs per exon and species; percentage of available sequence in relation to full length sequence is shown; * = exon sequence not full length; ––– = no data available

<table>
<thead>
<tr>
<th>Exon</th>
<th>Mus</th>
<th>Microcebus</th>
<th>Daubentonia</th>
<th>Otolemur</th>
<th>Tarsius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>236</td>
<td>239</td>
<td>239</td>
<td>239</td>
<td>242</td>
</tr>
<tr>
<td>2</td>
<td>292</td>
<td>292</td>
<td>274*</td>
<td>292</td>
<td>292</td>
</tr>
<tr>
<td>3</td>
<td>873</td>
<td>873</td>
<td>471*</td>
<td>873</td>
<td>640*</td>
</tr>
<tr>
<td>4</td>
<td>231</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>6</td>
<td>983</td>
<td>983</td>
<td>983</td>
<td>983</td>
<td>983</td>
</tr>
<tr>
<td>Σ</td>
<td>2739</td>
<td>2739</td>
<td>2319</td>
<td>2739</td>
<td>2281</td>
</tr>
<tr>
<td>%</td>
<td>100%</td>
<td>100%</td>
<td>84.6%</td>
<td>100%</td>
<td>83.2%</td>
</tr>
</tbody>
</table>

App. 3.2 B: Exons 1 to 6 of the VN2R2 gene in four primate species and the mouse (Vmn2r56) as reference; numbers show base pairs per exon and species; percentage of available sequence in relation to full length sequence is shown; * = exon sequence not full length; ps = exon sequence pseudogenised; ––– = no data available

<table>
<thead>
<tr>
<th>Exon</th>
<th>Mus</th>
<th>Microcebus</th>
<th>Daubentonia</th>
<th>Otolemur</th>
<th>Tarsius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>285</td>
<td>285</td>
<td>285</td>
<td>285</td>
<td>144</td>
</tr>
<tr>
<td>2</td>
<td>798</td>
<td>816</td>
<td>816</td>
<td>798</td>
<td>810ps</td>
</tr>
<tr>
<td>3</td>
<td>191</td>
<td>191</td>
<td>191</td>
<td>191</td>
<td>191</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>22</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>905</td>
<td>980</td>
<td>511*</td>
<td>890</td>
<td>—</td>
</tr>
<tr>
<td>Σ</td>
<td>2340</td>
<td>2418</td>
<td>1927</td>
<td>2288</td>
<td>1145</td>
</tr>
<tr>
<td>%</td>
<td>100%</td>
<td>100%</td>
<td>81.7%</td>
<td>98.4%</td>
<td>51.8%</td>
</tr>
</tbody>
</table>
### App. 3.3: Genbank accession numbers for whole-genome shotgun reads containing exonic data of VN2R1 and VN2R2 in primates (number of included exons in brackets)

<table>
<thead>
<tr>
<th>Species</th>
<th>VN2R1</th>
<th>VN2R2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Callithrix jacchus</em></td>
<td>ACFV01118612 (5-6)</td>
<td></td>
</tr>
<tr>
<td><em>Daubentonia madagascariensis</em></td>
<td>AGTM011617569 (1)</td>
<td>AGTM011790748 (1-2)</td>
</tr>
<tr>
<td></td>
<td>AGTM011566439 (2)</td>
<td>AGTM010361638 (2-3)</td>
</tr>
<tr>
<td></td>
<td>AGTM010374033 (3)</td>
<td>AGTM011731638 (5-6)</td>
</tr>
<tr>
<td></td>
<td>AGTM011537879 (4-6)</td>
<td></td>
</tr>
<tr>
<td><em>Gorilla gorilla</em></td>
<td>CABD02256199 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CABD02256200 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Homo sapiens</em></td>
<td>ADDF013464759 (5)</td>
<td>AEKP01177721 (6)</td>
</tr>
<tr>
<td></td>
<td>ADDF013464755 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Macaca fascicularis</em></td>
<td>CAEC01382845 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAEC01382844 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Macaca mulatta</em></td>
<td>AANU01195524 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AANU01195523 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Microcebus murinus</em></td>
<td>ABDC01032342 (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABDC01032343 (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABDC01032345 (5-6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABDC01360584 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABDC01360583 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Nomascus leucogenys</em></td>
<td>ADFV01169519 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ADFV01169520 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Otolemur garnetti</em></td>
<td>AAQR03107321 (1-3)</td>
<td>AAQR03119359 (1-2)*</td>
</tr>
<tr>
<td></td>
<td>AAQR03107323 (4)</td>
<td>AAQR03119358 (3+5)*</td>
</tr>
<tr>
<td></td>
<td>AAQR03107325 (5-6)</td>
<td>AAQR03119357 (6)*</td>
</tr>
<tr>
<td><em>Pan paniscus</em></td>
<td>AJFE01107863 (5-6)</td>
<td>AJFE01056479 (6)</td>
</tr>
<tr>
<td><em>Pan troglodytes</em></td>
<td>AACZ03021937 (5)</td>
<td>AACZ03117740 (6)</td>
</tr>
<tr>
<td><em>Papio anubis</em></td>
<td>AHZZ01103269 (5-6)</td>
<td></td>
</tr>
<tr>
<td><em>Pongo abelii</em></td>
<td>ABGA01384143 (5-6)</td>
<td>ABGA01258385 (6)</td>
</tr>
<tr>
<td><em>Saimiri boliviensis</em></td>
<td>AGCE01030805 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AGCE01030804 (6)</td>
<td></td>
</tr>
<tr>
<td><em>Tarsius syrichta</em></td>
<td>ABRT010347366 (1-2)</td>
<td>ABRT011131992 (1)</td>
</tr>
<tr>
<td></td>
<td>ABRT010347367 (3)</td>
<td>ABRT010517214 (2)</td>
</tr>
<tr>
<td></td>
<td>ABRT010394817 (5-6)</td>
<td>ABRT010517215 (3)</td>
</tr>
</tbody>
</table>

*a* = this information has been corrected in this thesis and deviates from the published manuscript

### App. 3.4: PAML results for VN2R1 and VN2R2; $\omega = d_{s}/d_{S}$, $p =$ proportion of sites in one class, $a =$ Likelihood ratio statistics (df = 2), $* = p<0.05$

<table>
<thead>
<tr>
<th>Locus</th>
<th>Class of sites</th>
<th>M0</th>
<th>M1a</th>
<th>M2a</th>
<th>M1a vs. M2a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\omega$ $p$</td>
<td>$\omega$ $p$</td>
<td>$\omega$ $p$</td>
<td>$\omega$ $p$</td>
<td></td>
</tr>
<tr>
<td>VN2R1</td>
<td>1 0.308</td>
<td>76.6% 0.135</td>
<td>76.6% 0.135</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 23.4%</td>
<td>1.000 8.4%</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 14.9%</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VN2R2</td>
<td>1 0.297</td>
<td>71.8% 0.097</td>
<td>76.4% 0.134</td>
<td>6.9*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 28.2%</td>
<td>1.000 22.2%</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 1.4%</td>
<td>22.818</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Functional promiscuity in a mammalian chemosensory system: Extensive expression of vomeronasal receptors in the main olfactory epithelium of mouse lemurs

Submitted in 2014 to Chemical Senses
by Philipp Hohenbrink$^{1,2}$, Silke Dempewolf$^1$, Elke Zimmermann$^1$, Nicholas I. Mundy$^{2,*}$ and Ute Radespiel$^{1,*}$

$^1$Institute of Zoology, University of Veterinary Medicine Hannover, Germany
$^2$Department of Zoology, University of Cambridge, Cambridge, UK
*These two authors contributed equally to the work

The vomeronasal organ (VNO) is functional in most terrestrial mammals, though progressively reduced in the primate lineage, and is used for intraspecific communication and predator recognition. Vomeronasal receptor (VR) genes comprise two families of chemosensory genes ($VIR$ and $V2R$) that have been considered to be specific for the VNO, but there is little knowledge of the expression of these genes outside of rodents. To explore the function of VR genes in mammalian evolution, we analysed and compared the expression of 64 $VIR$ and 2 $V2R$ genes in the VNO and the main olfactory epithelium (MOE) of the grey mouse lemur ($Microcebus murinus$), the primate with the largest known VR repertoire. We furthermore compared expression patterns in adults of both sexes and seasons, and in an infant.

A large proportion (83% – 97%) of the VR loci was expressed in the VNO of all individuals. The repertoire in the infant was as rich as in adults, indicating reliance on olfactory communication from early postnatal development onwards. Unexpectedly, we also detected extensive expression of VRs in the MOE, with proportions of expressed loci in individuals ranging from 29% to 45%. $TRPC2$, which encodes a channel protein crucial for signal transduction via VRs, was co-expressed in the MOE in all individuals indicating likely functionality of expressed VR genes in the MOE.

In summary, the large VR repertoire in mouse lemurs seems to be highly functional and we document unprecedented expression of VRs in the MOE, implying increased functionality of pheromone/kairomone detection in these primates. Given the differences in
the neural pathways of MOE and VNO signals, which project to higher cortical brain centres or the limbic system, respectively, this raises the intriguing possibility that the evolution of MOE-expression of VRs enabled mouse lemurs to adaptively diversify the processing of VR-encoded olfactory information.

4.1 Introduction

Most terrestrial mammals use two olfactory systems, the main olfactory system based on the main olfactory epithelium (MOE) and the accessory olfactory system based on the vomeronasal organ (VNO). The MOE has traditionally been considered to detect small odorants and a few pheromones (Restrepo et al. 2004), whereas the VNO is specialised for the detection of pheromones, signature mixtures and kairomones (e.g., predator cues) (Keverne 1999; Isogai et al. 2011) and is therefore essential for intraspecific communication and predator avoidance. Two types of vomeronasal receptors (VRs) are described which have been thought to be primarily expressed in the VNO based on studies in rodents: vomeronasal 1 receptors ($V_1R$s, Dulac and Axel 1995) and vomeronasal 2 receptors ($V_2R$s, Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). The two receptor types seem to be specialised for ligands of different size: $V_1R$s bind smaller molecules (Leinders-Zufall et al. 2000), whereas $V_2R$s bind larger peptides (like MHC: Leinders-Zufall et al. 2004; ESP1: Kimoto et al. 2005; and MUP: Chamero et al. 2007). The broader pattern of expression is yet unclear and a few cells of the MOE in mice and goats express single $V_1R$s (Wakabayashi et al. 2002; Karunadasa et al. 2006; Ohara et al. 2009). Both types of VRs use the cation channel $TRPC2$ in the signal transduction pathway (Liman et al. 1999). Although one study did not find $TRPC2$ expression outside of the mouse VNO (Zhang et al. 2010), another study reported a very faint signal for $TRPC2$ in the MOE on Northern blots and a faint signal in a small population of MOE cells by in situ hybridisation (Liman et al. 1999).

In primates it was observed that the size of the VNO correlates with its functionality: it is well developed and functional in strepsirrhine primates (lemurs and lorises), smaller but still functional in tarsiers and most New World monkeys and vestigial or absent in catarrhine primates (Old World monkeys, apes and humans) (Martin 1990; Smith et al. 2011). $V_1R$ repertoires in strepsirrhine primates (78-214 estimated genes) are as large as in rodents (89-239 estimated genes, Young et al. 2010), whereas only 2 intact $V_2R$ genes have been described in strepsirrhines (Hohenbrink et al. 2013) in contrast to about 100 intact genes in rodents (Young and Trask 2007). On the other hand, catarrhine primates lack functional $V_2R$s and the few
intact V1Rs are not considered to be functional, since TRPC2 is a pseudogene in catarrhines (Liman and Innan 2003; Zhang and Webb 2003; Young and Trask 2007; Young et al. 2010).

Strepsirrhine primates are ideal to study the vomeronasal system in non-model species as they heavily rely on olfactory communication (Jolly 1966; Perret 1995; DelBarco-Trillo et al. 2011) and olfactory predator recognition has been described in mouse lemurs (Sündermann et al. 2008; Kappel et al. 2011). Although large genomic repertoires of vomeronasal receptors (VRs) have been identified in strepsirrhine primates (Young et al. 2010; Yoder et al. 2014), little is known about the expression patterns of the genes. Thus it is currently unclear whether the pattern of expression found in mice is representative of other mammals, which has important consequences for understanding the functional role of VRs.

To further our understanding of the role of VNO and MOE in the olfactory system of mammals and to shed light on the evolution of this sensory modality in primates, we analyse the expression patterns of VRs and TRPC2 in grey mouse lemurs (Microcebus murinus). Among strepsirrhine primates mouse lemurs are an ideal model system as their genome is available (although yet incomplete) and the organisation of their V1R and V2R repertoires has been analysed (Young et al. 2010; Hohenbrink et al. 2012, 2013). Notably, they possess one of the largest predicted V1R repertoire of any mammal, comprising at least nine monophyletic gene clusters that have mostly evolved under positive selection (Hohenbrink et al. 2012). For this study we analysed expression of a large proportion of V1R and V2R loci in the VNO and MOE of both sexes and both seasons (reproductive and non-reproductive season). We also studied a young infant to identify expression patterns at a very early developmental age. We ask whether a high proportion of the VR repertoire is expressed in the VNO, and if this varies with age, sex or season. We also ask whether VR expression in the MOE is more widespread than currently believed.

4.2 Materials and Methods

Tissue collection

Complete VNOs and MOE tissues were collected from freshly deceased grey mouse lemurs that were euthanized for veterinary reasons or died naturally. No animal was sacrificed for the purpose of this study. All individuals were housed in the breeding colony of the Institute of Zoology of the University of Veterinary Medicine in Hannover under seasonal light regimes (see Wrogemann and Zimmermann 2001 for details). The international and national guidelines for the care and housing of animals were followed namely the NRC Guide
for the Care and Use of Laboratory Animals, the European Directive 2010/63/EU on the protection of animals used for scientific purposes, and the German Animal Welfare Act (licensed by the Bezirksregierung Hannover, reference number AZ 33.9-42502-05-10A080, and by Ordnungsamt, Gewerbe- und Veterinärautobteilung, Landeshauptstadt Hannover, AZ 42500/1H). Animals used in this study died between March 2012 and April 2013 and were not visually impaired. We selected four different adults to detect seasonal and sex differences (see Table 4.1 for details). We also tested one female infant and removed tissue from the maxilloturbinals of one adult as a negative control. Maxilloturbinal tissue is non-sensory (no olfactory mucosa) but nasal tissue in close proximity to the vomeronasal organ (Smith and Rossie 2008). If maxilloturbinal tissue was not contaminated with VNO cells during the removal, we assume that MOE tissue was also not contaminated with VNO cells and potential expression patterns of VRs in MOE tissue shown in this study relate to actual expression in the MOE. We were not able to obtain a MOE sample for the female that died during the non-reproductive season due to problems during the dissection. All tissue samples were stored in RNAlater (Qiagen) at -80°C immediately after removal and RNA was extracted up to one year after storage using the RNeasy Micro Kit (Qiagen). The 260 nm / 280 nm absorbance ratios of all extracted RNA samples were close to 2.0 indicating high RNA purity (instead of 1.8 for DNA). Transcription into cDNA was performed with the Quantitect Reverse Transcription Kit (Qiagen) and N6 primer, according to manufacturer’s instructions. The extraction and transcription kits both contain steps by the manufacturer to eliminate genomic DNA by DNase treatment. The successful extraction of RNA and synthesis of cDNA was confirmed by amplifying an intron-spanning segment of ACTB (beta-actin; ACTB-Exon-4-fw CTG TGC TGT CCC TGT ACG C, ACTB-Exon-6-rv AGT CCG C CT AGA AGC ATT TG), for which intronless cDNA was shorter on an agarose gel than genomic control DNA. Amplification and sequencing of an intron-spanning fragment of TRPC2 was used to confirm its expression. Primers for TRPC2 were TRPC2-A2 (TGA GCC AGG ACT ATG GCT TT) and TRPC2-B (Talarico 2006: CAG GTT CCC ACA CCA GAT G), which bind to exon 3 and 4, respectively (PCR conditions below). No long intron-containing bands were found on the agarose gel after amplification of TRPC2 and ACTB and together with the 260/280 absorption ratio and the two steps of genomic DNA elimination during RNA extraction and reverse transcription we exclude any contamination of our samples with genomic DNA.
Table 4.1: Tissue samples with information about age and sex of animal and season during the time of death; M = male; F = female; R = Reproductive season; NR = Non-reproductive season; MOE = Main olfactory epithelium; VNO = Vomeronasal organ; MT = Maxillo-turbinal. Light conditions during housing were adjusted to the Northern hemisphere.

<table>
<thead>
<tr>
<th>Animal</th>
<th>Sex</th>
<th>Season</th>
<th>Death</th>
<th>Age</th>
<th>Organ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zambo</td>
<td>M</td>
<td>R</td>
<td>2012-04-26</td>
<td>11.9 yrs</td>
<td>VNO, MOE</td>
</tr>
<tr>
<td>Uma</td>
<td>F</td>
<td>R</td>
<td>2013-03-14</td>
<td>7.9 yrs</td>
<td>VNO, MOE</td>
</tr>
<tr>
<td>Vincent</td>
<td>M</td>
<td>NR</td>
<td>2012-10-09</td>
<td>8.3 yrs</td>
<td>VNO, MOE</td>
</tr>
<tr>
<td>Tanja</td>
<td>F</td>
<td>NR</td>
<td>2012-09-20</td>
<td>6.3 yrs</td>
<td>VNO</td>
</tr>
<tr>
<td>Infant</td>
<td>F</td>
<td>-</td>
<td>2013-04-25</td>
<td>10 d</td>
<td>VNO, MOE</td>
</tr>
<tr>
<td>Ursina</td>
<td>F</td>
<td>R</td>
<td>2012-03-13</td>
<td>8.9 yrs</td>
<td>MT</td>
</tr>
</tbody>
</table>

Confirmation of transcribed loci

We developed locus-specific PCR assays to assess the expression of VRs as currently described. From the 107 previously published V1R sequences of the grey mouse lemur (Young et al. 2010) we excluded 29 loci, because they had no full length sequence (VN1R Mmur073) or did not fulfil the following criteria: 1) locus has at least 1% nucleotide differences to other loci (or else one of the two highly similar loci was excluded), 2) primers had not been rejected by the online software Primer3Plus (Untergasser et al. 2007) due to low/high annealing temperature or self assembly using default settings, and 3) the primer pair amplifies at least 500 bp of the target locus. Consequently, we designed locus-specific primer pairs for each of the remaining 78 V1R loci. The lowest sequence divergence between two loci for which locus-specific primer pairs were designed was ~3% (= 28 bp differences) and amplicons were at least 540 bp long. V1R loci are named VN1R Mmur000 to VN1R Mmur103 (882-1008 bp, Hohenbrink et al. 2012); V2R loci are named VN2R1 (2739 bp) and VN2R2 (2418 bp) (Hohenbrink et al. 2013).

The specificity of the newly designed V1R primer pairs was tested by amplifying and sequencing one cDNA sample (VNO tissue, male, reproductive season) and genomic DNA as a positive control. Sequencing genomic DNA enabled confirmation of specificity of each primer pair even if the locus is not expressed in the VNO sample. V2R primers were already validated (Hohenbrink et al. 2013). For the positive control genomic DNA was extracted from ear tissue using a DNeasy Tissue Kit (Qiagen) and a REPLI-g WGA kit (Qiagen). All VR fragments were amplified with MyTaq DNA polymerase (Bioline; 25 µl total volume containing 5 µl MyTaq Reaction Buffer, 1 µl of each primer [10 µM stock concentration], 0.1 µl Taq DNA polymerase [5 U/µl] and 1 µl of DNA) with the following PCR conditions: 94°C for 2 min, 40 times (94°C for 30 s, 60°C for 45 s, 72°C for 90 s), 72°C for 5 min. PCR products were sequenced on both strands using BigDye Terminator 3.1 (Applied Biosystems).
under standard conditions and run on an Applied Biosystems 3500 capillary sequencing machine. Consensus sequences of single genes were built with SeqMan 5.05 (DNASTAR Inc., Madison, WI, USA). Sequences were aligned and analysed using MEGA 5 (Tamura et al. 2011). The alignment of the sequenced fragments with the targeted loci showed that 64 of the 78 primer pairs were truly locus-specific. These V1R primer pairs and the V2R primers were used in subsequent PCRs under the same conditions with all tissue samples and run on 1% agarose gel. The presence and absence of bands was used to confirm or exclude the expression of loci in all tissue samples. All negative PCRs were repeated once to counteract PCR stochasticity and results are presented cumulatively. This method was preferred over next generation sequencing (RNA-Seq) because of the potential difficulty of correctly assembling RNA-Seq data for the large V1R gene family, with many closely related sequences.

We were able to design primers for loci of each monophyletic V1R cluster known from grey mouse lemurs (Hohenbrink et al. 2012). Four clusters were covered 100% (cluster II, IV, VI and VII). Also clusters III (75%) and V (71%) showed high coverage. However, for cluster VIII only one locus (20%) could be used, because the other four loci of the cluster were highly similar and designed primers turned out to be not locus-specific. Cluster I (35%; also known as V1Rstrep, Yoder et al. 2014) and IX (50%) are the two largest clusters in mouse lemurs and seemed to have evolved from rapid gene duplication. Cluster I was particularly difficult for primer design because of low nucleotide divergence: seven of nine loci that were excluded because of divergence below 1% belonged to this cluster.

4.3 Results

TRPC2 was expressed in all VNO and MOE tissue samples in the grey mouse lemur (Fig. 4.1). The concentration of the PCR products was slightly higher for all VNO samples than for MOE samples. There was no amplification of TRPC2 from the negative control of maxilloturbinal tissue, but all samples were positive for ACTB (not shown), indicating successful RNA extraction and cDNA synthesis. Each of the 64 V1R and the 2 V2R loci were expressed in at least one tissue sample. The specific expression patterns of VR loci in VNO and MOE tissues are shown in Table 4.2. The sample of maxilloturbinal tissue was negative for VR loci.
The proportion of expressed V1R loci ranged from 83% to 97% in the VNO and 28% to 45% in the MOE samples. The number of expressed genes was significantly higher in the VNO than in the MOE (observed vs. expected: $\chi^2 > 9.66$, df = 1, $p < 0.002$ in all four pairwise comparisons). We found no global differences in the number of expressed V1Rs between the two sexes or the two seasons neither in the VNO nor in the MOE, but one locus (VN1R Mmur074) was exclusively expressed in females. There were also no large scale differences between the adult individuals and the infant. Nevertheless, three loci (VN1R Mmur021, 036 and 045) were expressed in the VNO samples of all tested adults but were absent in the VNO sample of the infant. In contrast, VN1R Mmur053 was absent in the VNO tissue of all tested adults but was expressed in the infant.

At least one locus of each of the nine known monophyletic V1R clusters was expressed in each of the five VNO tissue samples. Four loci (VN1R Mmur004, 041, 044 and 059) were expressed in all nine tissue samples, whereas 25 V1R loci (39%) were exclusively expressed in the VNO. Fourteen of these V1R loci were expressed in all five tested VNO samples. In total, 48 V1R loci (75%) were expressed in all five VNO tissue samples meaning that variation in expression patterns in the VNO can only be found in 25% of the loci. More variation was found in the MOE samples. With the exception of three loci (VN1R Mmur030, 053 and 076) every locus that was expressed in the MOE was also expressed in the corresponding VNO sample of the same individual. The two V2R loci were expressed in the VNO of all individuals and the MOE of both females (including the infant) were positive for VN2R1.
Table 4.2: Expression of vomeronasal receptors in vomeronasal organ (VNO) or main olfactory epithelium (MOE) tissue. Cluster names (I-IX) are given for V1R loci (or uncl. = unclustered). Expressed loci are shown in grey. The total number of expressed loci per tissue sample is shown; R = Reproductive season, NR = Non-reproductive season, Inf = Infant, M = Male, F = Female.

| Tissue | Season | Sex | Mmur001 | Mmur002 | Mmur003 | Mmur004 | Mmur005 | Mmur006 | Mmur007 | Mmur008 | Mmur009 | Mmur010 | Mmur011 | Mmur012 | Mmur013 | Mmur014 | Mmur015 | Mmur016 | Mmur017 | Mmur018 | Mmur019 | Mmur020 | Mmur021 | Mmur022 | Mmur023 | Mmur024 | Mmur025 | Mmur026 | Mmur027 | Mmur028 | Mmur029 | Mmur030 | Mmur031 | Mmur032 | Mmur033 | Mmur034 | Mmur035 | Mmur036 | Mmur037 | Mmur038 | Mmur039 | Mmur040 | Mmur041 | Mmur042 | Mmur043 | Mmur044 | Mmur045 | Mmur046 | Mmur047 | Mmur048 | Mmur049 | Mmur050 | Mmur051 | Mmur052 | Mmur053 | Mmur054 | Mmur055 | Mmur056 | Mmur057 | Mmur058 | Mmur059 | Mmur060 | Mmur061 | Mmur062 | Mmur063 | Mmur064 | Mmur065 | Mmur066 | Mmur067 | Mmur068 | Mmur069 | Mmur070 | Mmur071 | Mmur072 | Mmur073 | Mmur074 | Mmur075 | Mmur076 | Mmur077 | Mmur078 | Mmur079 | Mmur080 | Mmur081 | Mmur082 | Mmur083 | Mmur084 | Mmur085 | Mmur086 | Mmur087 | Mmur088 | Mmur089 | Mmur090 | Mmur091 | Mmur092 | Mmur093 | Mmur094 | Mmur095 | Mmur096 | Mmur097 | Mmur098 | Mmur099 | Mmur100 | Mmur101 | Mmur102 | Mmur103 | Mmur104 | Mmur105 | Mmur106 |
| VNO    | R      | M   | Mmur001 | Mmur002 | Mmur003 | Mmur004 | Mmur005 | Mmur006 | Mmur007 | Mmur008 | Mmur009 | Mmur010 | Mmur011 | Mmur012 | Mmur013 | Mmur014 | Mmur015 | Mmur016 | Mmur017 | Mmur018 | Mmur019 | Mmur020 | Mmur021 | Mmur022 | Mmur023 | Mmur024 | Mmur025 | Mmur026 | Mmur027 | Mmur028 | Mmur029 | Mmur030 | Mmur031 | Mmur032 | Mmur033 | Mmur034 | Mmur035 | Mmur036 | Mmur037 | Mmur038 | Mmur039 | Mmur040 | Mmur041 | Mmur042 | Mmur043 | Mmur044 | Mmur045 | Mmur046 | Mmur047 | Mmur048 | Mmur049 | Mmur050 | Mmur051 | Mmur052 | Mmur053 | Mmur054 | Mmur055 | Mmur056 | Mmur057 | Mmur058 | Mmur059 | Mmur060 | Mmur061 | Mmur062 | Mmur063 | Mmur064 | Mmur065 | Mmur066 | Mmur067 | Mmur068 | Mmur069 | Mmur070 |
|        | F      |     |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
|        | Inf    | F   |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
| MOE    | R      | M   | Mmur001 | Mmur002 | Mmur003 | Mmur004 | Mmur005 | Mmur006 | Mmur007 | Mmur008 | Mmur009 | Mmur010 | Mmur011 | Mmur012 | Mmur013 | Mmur014 | Mmur015 | Mmur016 | Mmur017 | Mmur018 | Mmur019 | Mmur020 | Mmur021 | Mmur022 | Mmur023 | Mmur024 | Mmur025 | Mmur026 | Mmur027 | Mmur028 | Mmur029 | Mmur030 | Mmur031 | Mmur032 | Mmur033 | Mmur034 | Mmur035 | Mmur036 | Mmur037 | Mmur038 | Mmur039 | Mmur040 | Mmur041 | Mmur042 | Mmur043 | Mmur044 | Mmur045 | Mmur046 | Mmur047 | Mmur048 | Mmur049 | Mmur050 | Mmur051 | Mmur052 | Mmur053 | Mmur054 | Mmur055 | Mmur056 | Mmur057 | Mmur058 | Mmur059 | Mmur060 | Mmur061 | Mmur062 | Mmur063 | Mmur064 | Mmur065 | Mmur066 | Mmur067 | Mmur068 | Mmur069 | Mmur070 |
|        | F      |     |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
|        | Inf    | F   |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |

Primer pairs for the loci VN2R1 Mmur002, 005, 006, 014, 015, 016, 019, 029, 037, 062, 081, 092, 093 and 099 did not work or were not locus-specific. All non-listed loci from VN2R1 Mmur000 to 106 were excluded due to the three criteria (see 4.2 Material and Methods).
4.4 Discussion

We have demonstrated detailed expression patterns of vomeronasal receptors in a primate for the first time. We were able to show expression patterns for 64 of 78 loci (excluding potential alleles classified as loci) covering more than 80% of the V1R repertoire. As expected, the vast majority of the 64 tested V1Rs and both V2Rs were expressed in the VNO. However, unexpectedly, a substantial number of VRs (39 V1Rs and VN2R2) were also expressed in the MOE, which, given with TRPC2 co-expression in that organ, suggests that these VRs were likely functional. These results reveal a novel organisation of the chemosensory system and indicate greater functional overlap between the VNO and MOE than previously recognised.

The majority of the tested VRs (75%) were expressed in all five VNO samples indicating relatively low interindividual variation in the number of expressed VRs. Seasonal differences might have been expected, because some VRs might only be needed during the reproductive season, e.g., to find mating partners or to identify potentially oestrous females. However, we did not find a single locus that was solely expressed during the reproductive season in either olfactory organ, although these captive animals show the natural repertoire of seasonal behaviours (e.g. oestrus, mating) and seasonal variation in morphology (e.g. testis growth and body mass, Radespiel and Zimmermann 2001; Wrogemann et al. 2001).

Studies on mice showed that expression levels of VRs increased drastically from embryonic to 10-days postnatal age, but only changed marginally within the first seven months (Zhang et al. 2010). This was in contrast to expression of olfactory receptors in the MOE, where the same study found low expression levels at a lower age with constant increase of expression postnatally until the age of 3-4 months (already adult) with a slight decrease at a later age. This suggests that VRs unlike olfactory receptors need to be highly expressed from early life onwards. This is in concordance with our results in mouse lemurs where the infant showed a high proportion of expressed VRs similar to adult individuals (adulthood in mouse lemurs reached after less than 1 year, Glatston 1979). Infant mouse lemurs open their eyes at the age of four to six days, acoustically communicate with their mothers (Glatston 1979; Scheumann et al. 2007b), and start to leave the nest by the age of ~21 days (Lutermann 2001). Our results imply the potential for the use of olfactory signals between infants and nest mates (including mothers and siblings) or the early perception of predator cues.

In our study one male was almost twelve years old when he died. Wild mouse lemurs suffer from high predation (Scheumann et al. 2007a) and a near-complete population turnover
has been reported after five years (Kappeler and Rasolarison 2003). In captivity mouse lemurs in a normal photoperiodic regime are considered to be aged beyond the age of 7.5 years (E Zimmermann and U Radespiel, unpublished results) and can show signs of senescence. Following this definition three of four tested adult individuals were aged. However, aged mouse lemurs did not perform significantly worse than young animals in olfactory discrimination tasks, and only a minority of aged individuals showed altered behaviour during an olfactory reversal learning task (Joly et al. 2006). Given the high proportion of expressed genes in all our individuals, we do not predict any decreased functionality of the olfactory organs due to senescence, although deteriorating central nervous processing would be possible in analogy to the hearing system (Schopf 2013). Another explanation for an ongoing importance of the olfactory sense in aged animals could be the need to compensate an age-dependent decline of other senses (e.g., vision) with olfaction. However, although age-related visual impairments like cataracts and blindness have been described in captive grey mouse lemurs (Beltran et al. 2007), the individuals in our study were not visually impaired. A study on mice showed evidence of seven age-dependent expression profiles for VRs (Zhang et al. 2010), but testing only one infant and older adults did not allow to detect such profiles in our study. Mouse lemurs were not sacrificed for the purpose of this study and available samples were therefore limited. Animals would have to be euthanised systematically at different ages and in large numbers to collect information about age-dependent expression which cannot be supported for these primates from an ethical point of view.

We only found sex differences in expression in the VNO at a single locus. We assume that a large number of VRs are used to detect predator cues or signature mixtures to identify individuals and these types of information are equally important for both sexes, e.g., to minimize the predation risk or avoid inbreeding. Pheromones used for intraspecific communication should be equally relevant for males and females in most cases. Touhara & Vosshall (2009) assumed that male and female mice have the same set of pheromone receptors and that both sexes might show behavioural differences because of sex-specific neural circuits in the brain. Sex-specific signal transmission has been reported in mice where the same ligand and V2R receptor pair induces different behaviours in males and females (Haga et al. 2010; but also Halem et al. 1999). The VNO was also reported to be larger in male rats than in females (Segovia and Guillamón 1982), but no sex differences in VR expression are known. It has to be mentioned, though, that differences found between the VNO samples could also indicate individual differences. V1Rs show monoallelic expression
(Rodriguez et al. 1999; Roppolo et al. 2007) and the high number of expressed and most likely functional loci highlights the important role of the vomeronasal system for the sensory ecology of mouse lemurs. The function of each VR is still unclear and has to be analysed in further studies.

In the present study a large proportion of VRs (59%) as well as TRPC2 were also expressed in at least one MOE sample. This result was specific to MOE and VNO, since no VR or TRPC2 expression were detected in adjacent maxilloturbinal tissue. TRPC2 is essential for the functionality of the VNO, and is required for signal transduction of both V1Rs and V2Rs (Liman et al. 1999). Male mutant mice lacking TRPC2 did not attack intruding males and indiscriminately mounted males and females (Leypold et al. 2002; Stowers et al. 2002).

In our study both the VNO and MOE of mouse lemurs did express TRPC2 indicating functionality of the expressed VR genes in these chemosensory organs. In contrast, whereas in mice expression of TRPC2 in the VNO is similarly strong, only weak expression of TRPC2 has been so far reported in the MOE (Liman et al. 1999). The relatively strong expression of TRPC2 in the mouse lemur MOE was therefore unexpected, but in concordance with the high expression of VRs in this organ. Expression of single V1Rs in the MOE was reported in goats, mice and humans (Rodriguez et al. 2000; Wakabayashi et al. 2002; Karunadasa et al. 2006; Ohara et al. 2009). It was also shown in mice that 2-heptanone, a pheromone that binds the V1rb2 (Boschat et al. 2002), elicited strong signals in both the main and the accessory olfactory bulb (Xu et al. 2005), the brain structures that receive projections from the MOE or VNO, respectively (see Munger et al. 2009). Here we have found expression of far more VRs in the MOE of mouse lemurs than previously described in any mammal, including many loci expressed both in the MOE and VNO. The results were strengthened by the expression of TRPC2 in the MOE. Three main hypotheses may explain the involvement of the MOE in VR-mediated chemosensory pathways: 1) different ligand sensitivities in the two organs, 2) better coordination between the two organs or 3) different downstream neural pathways of the two organs.

The first hypothesis incorporates structural differences between MOE and VNO that facilitate the intake of volatile or non-volatile molecules, respectively. This argument is supported by the finding that in mouse lemurs the volatile phase of urine activates the MOE but not the VNO, which is only stimulated by urine in the liquid phase (Schilling et al. 1990). As the VNO lumen is filled with fluid that can be set in motion by vomeronasal pumps in the organ (Meredith et al. 1980; Meredith 1994), volatile molecules reach the vomeronasal
sensory epithelium less easily than non-volatile. Therefore, pheromones or other ligands bound by VRs being more or less volatile might be better perceived by one or the other olfactory organ. However, recent studies reject this classical view as both organs can perceive volatile as well as non-volatile pheromones (reviewed in Zufall and Leinders-Zufall 2007). Moreover, it would remain unclear, why the same VRs are expressed in both organs.

The other two hypotheses may explain the simultaneous expression: Ohara and colleagues (2009) suggested that V1Rs expressed in the MOE of goats might first detect pheromones that induce the flehmen response (animal raises the head and curls back the upper lip to facilitate the inflow of molecules to the VNO). According to the authors the pheromones are then quantitatively analysed with the V1Rs in the VNO and levels of airborne pheromones may be too subtle to be detected without the coordination of both olfactory organs.

However, as flehmen is not described in mouse lemurs (but found in ring-tailed lemurs, Bailey 1978), we present a third hypothesis: The expression of VRs in the MOE may be explained by the use of different neural pathways in both olfactory systems processing the signals in various brain regions (Mestre et al. 1992; Meisami and Bhatnagar 1998; Meredith 1998 and see review in Dulac and Wagner 2006; Touhara and Vosshall 2009). The MOE projects to the main olfactory bulb that after accessing paleocortical nuclei is connected to higher brain centres (Lledo et al. 2005) allowing adaptive responses based on experience. In contrast, the VNO bypasses cortical structures and projects directly to nuclei of the limbic system which mediates innate responses (Meisami and Bhatnagar 1998; von Campenhausen and Mori 2000; Dulac and Wagner 2006). Given the potential complexity in olfactory signal composition and transmission in a small nocturnal solitary forager such as the mouse lemur and the fundamental structural separation of the two olfactory organs including different neural pathways, it is likely that evolution may have favoured some degree of redundancy in the involved cells, receptors and structures to improve sensory abilities. For example, the availability of VR receptors in the MOE would allow mouse lemurs to better process the various olfactory signals that are produced and deposited in several species-specific marking behaviours (e.g. anogenital marking, head rubbing, urine washing, Glatston 1979; Buesching et al. 1998; Braune et al. 2005).

Future studies are needed to identify the full genomic repertoire of V1Rs and V2Rs based on the complete genome of mouse lemurs which has been sequenced at >160x coverage by J. Rogers and colleagues at the Baylor College of Medicine Human Genome Sequencing
Center and is currently in the assembly and annotation phase (A.D. Yoder, pers. comm.). We hypothesise that the simultaneous expression of a large number of the same VRs in MOE and VNO – which has never been shown in any mammalian species before – has evolved in mouse lemurs to adequately process a variety of complex olfactory signals, as separate neural pathways of both olfactory systems project to different brain regions performing special functions. Our results indicate a further blurring in the long presumed functional distinction between the VNO and MOE, following on from demonstration of pheromone detection in the MOE of some species, in some cases by olfactory receptors (Hudson and Distel 1986; Swann et al. 2001; Charra et al. 2012). More emphasis is needed on comparative adaptive function of these VRs in the MOE and VNO of mouse lemurs and other species with large VR repertoires and highly developed olfactory sense.

We thank the University of Veterinary Medicine Hannover for supporting the management of the mouse lemur colony and the Institute of Animal Breeding and Genetics for their technical help. We are also grateful to E. B. Keverne for helpful comments on the manuscript and we thank A. D. Yoder for her personal communication about the progress of the mouse lemur genome sequencing.

4.5 References


5. Signals of a divergent evolutionary history in the vomeronasal receptor genes of two sympatric mouse lemur species

In preparation to be submitted to a scientific peer-reviewed journal for publication in its current or slightly revised form
by Philipp Hohenbrink\textsuperscript{1,2}, Nicholas I. Mundy\textsuperscript{2} and Ute Radespiel\textsuperscript{1}

\textsuperscript{1}Institute of Zoology, University of Veterinary Medicine Hannover, Germany
\textsuperscript{2}Department of Zoology, University of Cambridge, Cambridge, UK

The genetic diversity of populations is shaped by demographic processes (e.g., migration, genetic drift) and selection. Many population studies attempt to exclude selection processes by analysing non-coding regions and it is questioned whether demographic histories can be analysed using expressed functional loci. Most terrestrial mammals possess functional vomeronasal receptors that are used in intraspecific communication and predator recognition. The two types of vomeronasal receptors (VRs), \( V1R \)s and \( V2R \)s, bind smaller ligands or larger peptides, respectively. Among primates, grey mouse lemurs possess the largest VR repertoire and it has evolved under strong positive selection. The high selection pressure suggests strong allele diversities of VRs within populations, but the diversity of these receptors has never been assessed in any primate. For a better understanding of population dynamics the aim of this study is to analyse and compare the evolution, genetic diversity and demographic history of a subset of VR genes in two sympatric mouse lemur species (\textit{Microcebus murinus} and \textit{M. ravelobensis}) in northwestern Madagascar. A total of 40 mouse lemurs (20 of each species) from one study site in the Ankarafantsika National Park were sequenced at 15 different \( V1R \) loci and the two \( V2R \) loci. All chosen loci had been shown before to be expressed in the vomeronasal organ of \textit{M. murinus}. A comparative analysis of the respective mismatch distributions as well as demographic and spatial modelling was used to reconstruct the demographic history of both mouse lemur populations that had previously been developed for \textit{M. murinus} on the basis of a non-coding locus. The ratio of nonsynonymous and synonymous substitutions within and between species and the distribution of nonsynonymous sites in the \( V1R \) protein were used to analyse the recent selection pressures on \( V1R \)s and \( V2R \)s.
M. ravelobensis showed a higher genetic diversity as expressed by a higher number of haplotypes and nucleotide diversity than M. murinus. The mismatch distributions and the expansion models suggested a more recent colonisation of the habitat by M. murinus than by M. ravelobensis whose colonisation happened earlier. The selected V1R loci contained a significant excess of synonymous substitutions and the detected nonsynonymous sites were randomly distributed within the receptor protein rather than concentrated on potential ligand binding sites, for example. The results indicate purifying selection acting on the majority of the VR loci of both species. After diversifying and evolving under positive selection as the general mode of selection in the more distant past, the selection pressure in recent history seemed to have stabilised the function of the receptors. However, only a subset of the V1R repertoire was analysed and potentially only a small number of V1Rs is evolving adaptively within populations.

5.1 Introduction

Genetic diversity is shaped by mutations and changes in allele frequencies due to demographic processes (e.g., migration, genetic drift) or selection (Frankham et al. 2002, pp. 44, 178; Guschaniski 2004). The level of genetic diversity depends on the historical and current population size, natural selections, different mutation rates and demographic events like population bottlenecks and migration among populations (Frankham et al. 2002, p. 68) but also anthropogenic disturbances. Genetic diversity within populations is essential to react to environmental changes (long-term survival) and maintain reproductive fitness (short-term survival, Frankham et al. 2002, p. 23; Guschaniski 2004). Whereas small populations are more affected by genetic drift than they are by selection, larger populations are more influenced by selection, which acts upon the phenotype and correspondingly on the DNA (Frankham et al. 2002, p. 214). Demographic processes are typically investigated by employing neutral loci (non-coding regions) to exclude selection processes (e.g., Kaessmann et al. 1999; Schneider et al. 2010; Kawamoto et al. 2013). However, there are some notable exceptions (e.g., Excoffier 2002; Quach et al. 2013). We therefore ask if demographic processes can still be recognised even in functional loci.

In the present study, we exemplarily use functional loci of vomeronasal receptor (VR) genes that are expressed in the vomeronasal organ of most terrestrial mammals (see Keverne 1999). The two types of VRs – V1Rs (Dulac and Axel 1995) and V2Rs (Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997) – bind small volatile molecules
(Leinders-Zufall et al. 2000) or larger peptide (Leinders-Zufall et al. 2004; Kimoto et al. 2005; Chamero et al. 2007), respectively. Both are G-protein coupled chemosensory receptors with seven transmembrane regions and chemosensory receptors were reported to be one of the functional categories that are most frequently found to be under positive selection in a variety of animal taxa (Kosiol et al. 2008). Among primates, mouse lemurs (Microcebus spp.) have the largest known repertoire of VRs (Young et al. 2010; Hohenbrink et al. 2013). The V1R repertoire consists of 214 estimated intact genes and at least 2 V2R genes showing a strong bias towards V1Rs which is less distinct in rodents (mice: 239 V1Rs and 123 V2Rs, rats: 108 V1Rs and 87 V2Rs, numbers are estimated intact genes according to Young and Trask 2007; and Young et al. 2010). The organisation of the V1R repertoire has been previously described (Hohenbrink et al. 2012): it consists of nine monophyletic gene clusters and a few unclustered loci. Seven of nine clusters have evolved under strong positive selection. However, genetic diversity within populations or between species has not been assessed for these receptors yet. The level of diversity allows drawing conclusions about the evolution within the population or species. Narrow-tuned receptors are predicted to allow less variation, whereas rapidly evolving receptors should result in large variation. The phylogeography of many taxa has been influenced by glacial events, in particular during the Pleistocene (Hewitt 1999), and the genetic diversity can be used to identify recent expansions in mouse lemurs, which live endemically on Madagascar. The explosive speciation in this group has been tried to be explained by several hypotheses based on river systems and Pleistocene climate change on the island (Martin 1972; Wilmé et al. 2006; Olivieri et al. 2007).

In this study we analyse the genetic diversity of VRs within a wild population and in two sympatric species of mouse lemurs (M. murinus and M. ravelobensis) in order to identify the evolutionary and demographic history of both species. Here, the two species share the same habitat, predators and climate, but the colonisation histories are potentially different. It was shown before that signals of two recent spatial expansions are present in the mitochondrial D-loop of the M. murinus population analysed in our study (Schneider et al. 2010). With the analysis of a larger number of 15 V1Rs and 2 V2Rs we want to look on the most recent evolution of mouse lemurs and VRs and try to reproduce similar results based on functional loci. In particular, we run expansion models to reconstruct Pleistocene colonisation events in both species based on VRs. The pattern of nonsynonymous and synonymous substitutions is used to identify recent selection pressures on VRs.
5.2 Material and Methods

Data collection

We extracted DNA from 20 grey mouse lemurs, *M. murinus* (10 male, 10 female), and 20 golden-brown mouse lemurs, *M. ravelobensis* (10 male, 10 female, see App. 5.1 for details), using a phenol-chloroform protocol and a REPLI-g WGA kit (Qiagen). Small ear tissue samples were collected between May and October 2008 in the Ankarafantsika National Park in northwestern Madagascar by S. Thorén (authorisation no. 062/08/MEEFT/SG/DGEF/DSA/SSE). All animals were live-trapped in the 30.6 ha long-term study site JBA (46°48’E, 16°19’S; for details about the trapping procedure see Thorén et al. 2010), and we selected individuals from different trapping locations to minimise the risk of sampling several individuals from the same sleeping group who are most likely related (Radespiel et al. 2001, 2009). All sampled animals of the same species are defined to belong to the same population.

We designed locus-specific primer pairs for 15 *VIR* and both *V2R* loci using the online software Primer3Plus (Untergasser et al. 2007). All loci (see Table 5.1 for details) were expressed in the VNO of the grey mouse lemur (Hohenbrink et al., submitted, or chapter 4). We used twelve loci from seven of the nine monophyletic *VIR* clusters and three unclustered loci. Clusters I and VIII were not analysed because designed primer pairs did not work for the eight expressed loci of cluster I and the one expressed locus of cluster VIII. The PCR amplicons of intronless *VIR* loci covered the whole locus as all primers bind externally (i.e., outside of the coding sequence). *VN2R1* consists of six exons and because of long intron sequences between the exons we designed exon-specific external primers. In contrast to our first description of *V2Rs* (Hohenbrink et al. 2013) *VN2R2* consists of only five exons. Previously we had used cDNA of extracted RNA without introns and sequenced a fragment spanning exon 3 to 5 according to closely related *V2Rs* of family D in mice. A short fragment (22 bp) between these exons was assigned to exon 4. However, the present study which used external primers revealed that the 22 bp of exon 4 rather belonged to exon 3 which was longer in mouse lemurs than in mice. The analogous exon 4 of mice was missing in mouse lemurs, which was in concordance with genomic data of the two strepsirrhines *Daubentonia madagascariensis* and *Otolemur garnetti* where no “exon 4” had been detected (App. 3.2 B), but retrospectively a similar 22 bp insertion at the end of exon 3 was found. Consequently, exon 5 and 6 were the 4th and 5th exon, respectively. The reading frame was intact and this error had no further consequences for the results presented previously (Hohenbrink et al. 2013). In the present study we could not design external primers for exon 2 of *VN2R2* because
of missing genomic data. Due to internal primers we are missing 108 bp (= 13.2%) of the sequence of exon 2 and therefore results on VN2R2 are based on 95.7% of the whole coding sequence.

We used MyTaq DNA polymerase for amplification (Bioline; 25 µl total volume containing 5.0 µl MyTaq Reaction Buffer, 1 µl of each primer [10 µM stock concentration], 0.1 µl Taq DNA polymerase [5 U/µl] and 1 µl of DNA) with the following PCR conditions: 94°C for 2 min, 40 times (94°C for 30 s, 60°C for 45 s, 72°C for 90 s), 72°C for 5 min. PCR products were sequenced on both strands using BigDye Terminator 3.1 (Applied Biosystems) under standard conditions and run on an Applied Biosystems 3500 capillary sequencing machine. Consensus sequences were built with SeqMan 5.05 (DNASTAR Inc., Madison, WI, USA). Sequences were aligned and analysed using MEGA 5 (Tamura et al. 2011). Before data analyses we concatenated the exons of the V2Rs.

Table 5.1: V1R and V2R loci analysed (with corresponding gene cluster and total length); uncl = unclustered

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cluster</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>VN1R Mmur001</td>
<td>IV</td>
<td>909 bp</td>
</tr>
<tr>
<td>VN1R Mmur011</td>
<td>VI</td>
<td>930 bp</td>
</tr>
<tr>
<td>VN1R Mmur031</td>
<td>V</td>
<td>909 bp</td>
</tr>
<tr>
<td>VN1R Mmur033</td>
<td>uncl</td>
<td>942 bp</td>
</tr>
<tr>
<td>VN1R Mmur040</td>
<td>II</td>
<td>948 bp</td>
</tr>
<tr>
<td>VN1R Mmur041</td>
<td>V</td>
<td>906 bp</td>
</tr>
<tr>
<td>VN1R Mmur043</td>
<td>VI</td>
<td>1005 bp</td>
</tr>
<tr>
<td>VN1R Mmur048</td>
<td>VI</td>
<td>957 bp</td>
</tr>
<tr>
<td>VN1R Mmur049</td>
<td>VII</td>
<td>879 bp</td>
</tr>
<tr>
<td>VN1R Mmur060</td>
<td>V</td>
<td>906 bp</td>
</tr>
<tr>
<td>VN1R Mmur065</td>
<td>uncl</td>
<td>1008 bp</td>
</tr>
<tr>
<td>VN1R Mmur066</td>
<td>uncl</td>
<td>897 bp</td>
</tr>
<tr>
<td>VN1R Mmur067</td>
<td>III</td>
<td>921 bp</td>
</tr>
<tr>
<td>VN1R Mmur074</td>
<td>IX</td>
<td>918 bp</td>
</tr>
<tr>
<td>VN1R Mmur075</td>
<td>IV</td>
<td>909 bp</td>
</tr>
<tr>
<td>VN2R1</td>
<td>V2R</td>
<td>2739 bp</td>
</tr>
<tr>
<td>VN2R2</td>
<td>V2R</td>
<td>2310 of 2418 bp*</td>
</tr>
</tbody>
</table>

* = locus 3 bp longer in M. ravelobensis

Data analyses

For each locus we used DnaSP 5.10 (Librado and Rozas 2009) to unphase the two alleles of each diploid sequence and to identify the different haplotypes. Nucleotide diversity and haplotype diversity (expected heterozygosity, or gene diversity, Nei 1987) were calculated with DnaSP to measure the genetic variation within the population/species.
The software Arlequin 3.5 (Excoffier and Lischer 2010) was used to run two models (demographic and spatial expansion) on the mismatch distributions of the haplotypes to analyse demographic events. A mismatch distribution is a distribution of the number of nucleotide mismatches between all pairs of nucleotide sequences of one locus and one population. Consequently each individual (homozygous or heterozygous) always enters two sequences into the data pool. The shape of the distribution indicates demographic events like past expansions or genetic bottlenecks and mismatch distributions can be easily compared between loci or species in this study because of the same sample size of 40 nucleotide sequences per distribution. It is usually multimodal in populations at demographic equilibrium, but unimodal after a demographic (Rogers and Harpending 1992) or spatial expansion in a population (Ray et al. 2003; Excoffier 2004). Demographic expansions usually result from past genetic bottlenecks, whereas spatial expansions usually follow colonisations by founder individuals. We tested two models for the respective expansion to evaluate the evidence for a preceding colonisation event (Schneider and Excoffier 1999; Excoffier 2004). The null hypothesis assumes that the observed mismatch distribution does not differ from the distribution under one demographic or spatial expansion. The models also calculate $\tau$- (tau-) values that reflect the age of the expansion. Higher values indicate that the expansion happened further in the past. Arlequin was also used to calculate neutrality tests (Tajima’s D and Fu’s Fs) with 1000 simulated samples. Population contractions can result in significant positive values of Tajima’s D, whereas negative values indicate population expansions (Tajima 1989b, a). Fu’s Fs tests show negative values if the data contains an excess of rare haplotypes also indicating population expansions (Fu 1997). In combination positive D and positive Fs values indicate an excess of intermediate-frequency alleles after population subdivision or balancing selection, whereas negative D and negative Fs values reflect relative excess of rare variants and reveal population growth (Bamshad et al. 2002). Fu’s Fs values were considered significant at 5% significance level with a $p$-value below 0.02 rather than 0.05 as recommended in the Arlequin software manual (p. 143 of cited version).

DnaSP was also used to estimate the number of synonymous ($d_S$) and nonsynonymous substitutions ($d_N$) per site (Jukes and Cantor 1969; Nei and Gojobori 1986) and $d_N/d_S$ ratios were calculated. McDonald-Kreitman tests (= MKT, McDonald and Kreitman 1991) were conducted online (Egea et al. 2008) to calculate the ratio of fixed differences to polymorphic differences for synonymous and nonsynonymous sites. Here, all haplotypes of *M. murinus*
and the most closely related haplotype sequence of \textit{M. ravelobensis} were entered to test for positive selection in \textit{M. murinus}, and vice versa to test in \textit{M. ravelobensis}.

It was analysed in which region of the \textit{V1R} protein (transmembrane, extra- or intracellular region) the nonsynonymous sites of \textit{V1R} loci were located (for details on the method used see Hohenbrink et al. 2012). Observed vs. expected \(\chi^2\)-tests were used to compare the observed distribution of nonsynonymous sites in the \textit{V1R} protein with the expected distribution using Statistica 6.1 (StatSoft, Inc., Tulsa, OK). A previous study on \textit{V1Rs} in strepsirrhines reported that the ligand binding site of the \textit{V1R} protein is potentially formed by about half of the 4\textsuperscript{th} and 5\textsuperscript{th} transmembrane region and the in-between 2\textsuperscript{nd} extracellular loop (= 3\textsuperscript{rd} extracellular region) (Yoder et al. 2014). This estimation is based on data of cluster I, but assuming structural similarities between clusters, we tested if nonsynonymous substitutions were concentrated on the binding site proposed by Yoder et al. (2014). All statistical comparisons of dependent data between the two species were conducted with the Wilcoxon Matched Pairs test in Statistica. Here, the sample size was large enough to ignore \(p\)-value corrections that would have been necessary for smaller sample sizes (Mundry and Fischer 1998).

5.3 Results

\textit{Differences in genetic diversity}

For the \textit{V1R} loci we found 0 to 10 polymorphic sites per locus in \textit{M. murinus} (= 0.0\% - 1.04\% of total length) and 1 to 28 sites in \textit{M. ravelobensis} (= 0.11 – 3.12\%, Table 5.2), respectively. The \textit{M. murinus} population had on average 4.5 different haplotypes (ranging from 1 to 11) per locus, whereas the \textit{M. ravelobensis} population contained on average 9.1 different haplotypes (ranging from 2 to 17). No haplotypes were shared between species (minimum interspecific distance of haplotypes was 5 bp in \textit{VN1R Mmur040}; minimum interspecific distance on protein level was 1 amino acid in \textit{Mmur075}). The number of haplotypes was significantly higher in \textit{M. ravelobensis} \((n = 15, Z = 3.15, p = 0.001)\) with 14 out of 15 individual \textit{V1R} loci containing more haplotypes in this species than in \textit{M. murinus}. \textit{M. ravelobensis} also showed a significantly higher nucleotide diversity \((n = 15, Z = 2.78, p = 0.005)\), a significantly higher haplotype diversity \((n = 15, Z = 2.44, p = 0.015)\) and a significantly higher number of polymorphic sites \((n = 15, Z = 3.04, p = 0.002)\). At the level of amino acid sequences, in \textit{V1Rs} the number of unique amino acid sequences was decreased by about one third of the haplotypes in both species (Table 5.2). One notable exception was locus
VNIR Mmur075 that had 7 DNA-haplotypes in *M. murinus* that all translate to the same amino acid sequence, although its 9 DNA-haplotypes in *M. ravelobensis* translate to 6 different amino acid sequences.

V2R loci had more haplotypes in both species than the V1R loci (Table 5.2). However, the number of polymorphic sites in V2Rs compared to V1Rs was mainly increased in *M. murinus*. Here, one polymorphic site was found per 98 bp or 151 bp in VN2R1 or VN2R2, respectively, which was below the 25%-75%-quartile range of V1Rs: 152 – 314 bp (median 227). In contrast, the number of polymorphic sites was similar between V1Rs and V2Rs in *M. ravelobensis* (one polymorphic site per 80 bp in VN2R1 and 86 bp in VN2R2 compared to 70 bp median in V1Rs) and within the quartile range of 45 – 106 bp for V1Rs. The number of unique amino acid sequences decreased considerably in VN2R2, but there was almost no decrease in VN2R1 indicating that here most haplotypes differed by at least one nonsynonymous substitution. Also no haplotype was shared between species (minimum interspecific distance was 27 bp or 13 amino acids for the two V2Rs). The results showed that the genetic diversity in V1Rs and VN2R2 differs between the species, whereas it was equally high in VN2R1 of both species.

### Table 5.2: Number of haplotypes, number of different amino acid sequences (= diff. AA sequences), nucleotide diversity, haplotype diversity and number of polymorphic sites for each locus and species; Mmur = *M. murinus*; Mrav = *M. ravelobensis*, Ø = V1R mean

<table>
<thead>
<tr>
<th>Locus</th>
<th>No. haplotypes</th>
<th>No. diff. AA sequences</th>
<th>Nucleotide diversity</th>
<th>Haplotype diversity</th>
<th>No. of polymorphic sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mmur</td>
<td>Mrav</td>
<td>Mmur</td>
<td>Mrav</td>
<td>Mmur</td>
</tr>
<tr>
<td>001</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>8</td>
<td>.00073</td>
</tr>
<tr>
<td>011</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>.00194</td>
</tr>
<tr>
<td>031</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>.00141</td>
</tr>
<tr>
<td>033</td>
<td>3</td>
<td>9</td>
<td>2</td>
<td>9</td>
<td>.00089</td>
</tr>
<tr>
<td>040</td>
<td>6</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>.00074</td>
</tr>
<tr>
<td>041</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>.00141</td>
</tr>
<tr>
<td>043</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>.00048</td>
</tr>
<tr>
<td>048</td>
<td>11</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>.00170</td>
</tr>
<tr>
<td>049</td>
<td>2</td>
<td>9</td>
<td>2</td>
<td>8</td>
<td>.00033</td>
</tr>
<tr>
<td>060</td>
<td>3</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>.00016</td>
</tr>
<tr>
<td>065</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>.00000</td>
</tr>
<tr>
<td>066</td>
<td>4</td>
<td>17</td>
<td>3</td>
<td>14</td>
<td>.00139</td>
</tr>
<tr>
<td>067</td>
<td>2</td>
<td>14</td>
<td>1</td>
<td>9</td>
<td>.00015</td>
</tr>
<tr>
<td>074</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>10</td>
<td>.00093</td>
</tr>
<tr>
<td>075</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>6</td>
<td>.00098</td>
</tr>
<tr>
<td>Ø</td>
<td>4.5</td>
<td>9.1</td>
<td>3.0</td>
<td>6.6</td>
<td>.00088</td>
</tr>
<tr>
<td>VN2R1</td>
<td>25</td>
<td>28</td>
<td>24</td>
<td>25</td>
<td>.00332</td>
</tr>
<tr>
<td>VN2R2</td>
<td>14</td>
<td>27</td>
<td>8</td>
<td>20</td>
<td>.00053</td>
</tr>
</tbody>
</table>
Table 5.3: Results of the neutrality tests, \( \tau \)-values of demographic (\( \tau_{\text{demo}} \)) and spatial model (\( \tau_{\text{spat}} \)) and \( d_S/d_S \) ratios for each locus and species; red values of neutrality tests are significant at 5% significance level; \( \text{Mmur} = M. \text{murinus} \); \( \text{Mrav} = M. \text{ravelobensis} \), — = \( d_S \) was zero

<table>
<thead>
<tr>
<th>Locus</th>
<th>Tajima’s D</th>
<th>Fu’s Fs</th>
<th>( \tau_{\text{demo}} )</th>
<th>( \tau_{\text{spat}} )</th>
<th>( d_S/d_S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>-0.70</td>
<td>1.46</td>
<td>-3.83</td>
<td>0.78</td>
<td>6.28</td>
</tr>
<tr>
<td>011</td>
<td>2.18</td>
<td>-0.42</td>
<td>3.96</td>
<td>-0.70</td>
<td>5.02</td>
</tr>
<tr>
<td>031</td>
<td>-0.24</td>
<td>-0.48</td>
<td>0.36</td>
<td>-2.05</td>
<td>1.88</td>
</tr>
<tr>
<td>033</td>
<td>0.42</td>
<td>0.60</td>
<td>1.42</td>
<td>1.21</td>
<td>0.00</td>
</tr>
<tr>
<td>040</td>
<td>-1.02</td>
<td>0.76</td>
<td>-2.38</td>
<td>2.13</td>
<td>0.88</td>
</tr>
<tr>
<td>041</td>
<td>0.86</td>
<td>-0.66</td>
<td>1.37</td>
<td>-1.91</td>
<td>1.50</td>
</tr>
<tr>
<td>043</td>
<td>-1.25</td>
<td>-0.74</td>
<td>-1.25</td>
<td>-1.49</td>
<td>0.48</td>
</tr>
<tr>
<td>048</td>
<td>-1.13</td>
<td>-1.98</td>
<td>-4.48</td>
<td>-6.69</td>
<td>1.77</td>
</tr>
<tr>
<td>049</td>
<td>0.37</td>
<td>-0.70</td>
<td>0.84</td>
<td>-2.44</td>
<td>2.98</td>
</tr>
<tr>
<td>060</td>
<td>-1.30</td>
<td>-0.86</td>
<td>-2.03</td>
<td>-0.89</td>
<td>3.00</td>
</tr>
<tr>
<td>065</td>
<td>0.00</td>
<td>-0.56</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>066</td>
<td>0.16</td>
<td>-0.50</td>
<td>1.29</td>
<td>-2.34</td>
<td>2.82</td>
</tr>
<tr>
<td>067</td>
<td>-0.56</td>
<td>0.28</td>
<td>-0.22</td>
<td>-2.63</td>
<td>0.21</td>
</tr>
<tr>
<td>074</td>
<td>-1.56</td>
<td>-0.87</td>
<td>-1.77</td>
<td>-0.63</td>
<td>3.73</td>
</tr>
<tr>
<td>075</td>
<td>-1.27</td>
<td>-0.23</td>
<td>-2.69</td>
<td>1.61</td>
<td>1.00</td>
</tr>
<tr>
<td>VN2R1</td>
<td>1.15</td>
<td>-0.08</td>
<td>-6.78</td>
<td>-13.58</td>
<td>0.48</td>
</tr>
<tr>
<td>VN2R2</td>
<td>-1.39</td>
<td>-0.78</td>
<td>-6.33</td>
<td>-17.25</td>
<td>2.54</td>
</tr>
</tbody>
</table>

Divergent demographic histories

Single neutrality tests were significantly negative in four \( V1R \) loci in \( M. \text{murinus} \) (Table 5.3, Tajima’s D in \( \text{VN1R Mur074} \) and Fu’s Fs in \( \text{Mmur001, 048 and 060} \) and both neutrality tests were significantly negative for locus \( \text{Mmur048} \) in \( M. \text{ravelobensis} \). No significantly positive values were found and test values did not differ significantly between both species (Tajima’s D: \( n = 15, Z = 0, p = 1.000 \); Fu’s Fs: \( n = 15, Z = 0.68, p = 0.500 \)). \( \text{VN2R1} \) and \( \text{VN2R2} \) had significantly negative Fs values in both species (Table 5.3). In \( M. \text{murinus} \) 9 of 15 Tajima’s D and 8 of 15 Fu’s Fs were negative and in \( M. \text{ravelobensis} \) 11 of 15 D and 11 of 15 Fs were negative, but no generalised pattern could be observed.

The mismatch distributions showed huge variation between loci and species (Fig. 5.1 and App. 5.2). Most \( V1R \) loci (\( n = 11 \)) had half-bell shaped distributions close to zero pairwise differences in \( M. \text{murinus} \) or the peak was at zero (Fig. 5.1 A). Furthermore, \( M. \text{murinus} \) had four unimodal distributions (\( \text{VN1R Mur011, 033, 066 and 074} \) but no multimodal or ragged distributions. In contrast, three loci in \( M. \text{ravelobensis} \) showed ragged distributions (\( \text{VN1R Mur001, 040, and 067, see Fig. 5.1 B} \) and six mostly broad unimodal distributions (\( \text{VN1R Mur033, 041, 060, 066, 074 and 075} \)). The remaining \( V1R \)s showed half-bell shaped distributions similar to \( M. \text{murinus} \). Notably, unimodal distributions in \( M. \)
murinus were still close to zero pairwise differences with a high peak, whereas in M. ravelobensis the distributions were generally broader and flat (Fig. 5.1 C + D). The occurrence of half-bell shaped, unimodal and ragged distributions in both species is listed in Table 5.4. Although most loci (8 of 15) had similar types of distributions in both species, the distributions of most of the other loci showed higher variation in M. ravelobensis than in M. murinus. In V2Rs (App. 5.2), the mismatch distributions in M. murinus showed a ragged distribution for VN2R1 and a unimodal distribution for VN2R2. The distributions in M. ravelobensis were unimodal for both V2Rs.

Fig. 5.1: Observed and simulated mismatch distributions of VN1R Mmur001 and 066 in M. murinus (Mmur, left side, grey) and M. ravelobensis (Mrav, right side, orange); the two loci were selected to show the three observed types of distributions: half-bell shaped (A), unimodal (C + D) and ragged (B); mismatch distributions of the remaining loci are shown in App. 5.2; simulated (d) = simulated after demographic expansion model (line with circles), simulated (s) = simulated after spatial expansion model (line with crosses)
Table 5.4: Pairwise occurrence of observed mismatch distributions for the V1R loci in both species; grey = more variation within M. murinus; orange = more variation within M. ravelobensis; HB = Half-bell shape distribution; UM = Unimodal distribution; RG = Ragged distribution; Σ = Sum

<table>
<thead>
<tr>
<th></th>
<th>HB</th>
<th>UM</th>
<th>RG</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. murinus</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>M. ravelobensis</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Σ</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

The demographic expansion model was significant in two V1R loci of M. ravelobensis (VN1R Mmur041: \( p = 0.02 \); Mmur075: \( p = 0.02 \)). The spatial expansion model was significant in two V1R loci of M. murinus (Mmur040: \( p = 0.01 \); Mmur075: \( p = 0.04 \)). The expansions were rejected in these cases because the observed distribution differed significantly from the simulated distribution. In all other cases (including V2Rs) observed mismatch distributions were not significantly different from the expected distribution under one demographic or spatial expansion. The \( \tau \)-value was significantly larger in M. ravelobensis compared to M. murinus using both models (Table 5.3; demographic expansion: \( n = 15, Z = 2.38, p = 0.017 \); spatial expansion: \( n = 15, Z = 2.44, p = 0.015 \)) indicating an older beginning of an expansion in M. ravelobensis compared to M. murinus.

**Evidence for purifying selection**

The \( d_S \) of V1Rs was significantly higher than \( d_N \) in both species (M. murinus: \( n = 15, Z = 2.92, p = 0.004 \); M. ravelobensis: \( n = 15, Z = 2.39, p = 0.017 \)). The \( d_S/d_N \) ratios of V1Rs did not differ significantly between species (Table 5.3; \( n = 13, Z = 1.01, p = 0.311 \)). The MKT was only significant for VN1R Mmur066 in M. ravelobensis (\( \chi^2 = 4.89, p = 0.026 \)) indicating that the locus had a significant excess of nonsynonymous substitutions in M. ravelobensis. All other MKTs for V1Rs or V2Rs were not significant (\( p > 0.05 \)).

The distribution of nonsynonymous sites in the domains of the V1R protein did not differ significantly from the expected distribution in any species when looking at the data of all V1R loci combined (M. murinus: \( \chi^2 = 3.2, df = 2, p = 0.200 \); M. ravelobensis: \( \chi^2 = 1.0, df = 2, p = 0.606 \)). Actually, the observed distributions (extracellular sites – sites in transmembrane region – intracellular sites) in M. murinus (8–12–12) and M. ravelobensis (30–55–32) were
highly similar to the expected distributions of 9–15–8 and 34–55–28, respectively. Also, single loci did not show significant deviations from the expected distributions (for example, VNIR Mmur066, the only locus with significant MKT: $\chi^2 = 2.6$, df = 2, $p = 0.268$; all other loci also with $p > 0.05$). The binding site proposed by Yoder and colleagues (2014) contained about 20% of all amino acids of the V1R protein. In M. murinus 21.9% of the nonsynonymous substitutions were located in this binding site. In M. ravelobensis the proportion was 19.7% which does not reveal any excess of nonsynonymous substitutions in this part of the protein for the two species. Also, single loci did not show concentrations of nonsynonymous substitutions in this potential binding site.

5.4 Discussion

The results showed divergent evolution of the VRs in M. murinus and M. ravelobensis. These include differences in the genetic diversity and the demographic histories. However, purifying selection acting on most VRs was indicated in both species.

Demographic histories of two sympatric mouse lemur species

Three types of mismatch distributions were found: 1) Half-bell shaped distributions show a lack of variation which indicates purifying selection or recent bottlenecks. 2) Unimodal distributions are seen after one demographic or spatial expansion. 3) Ragged distributions are typical for populations at demographic equilibrium where colonisation events are very old. After enough generations without strong purifying selection and demographic events a half-bell shaped distribution will turn into a unimodal distribution and then potentially turn into a ragged distribution.

The mismatch distributions did not differ significantly from the simulated distribution after one demographic or spatial expansion in most loci of both species. Demographic expansions are seen in stationary populations, whereas spatial expansions increase the range of the population over time and space (Excoffier and Lischer 2010). Given the similar consequences of both colonisation scenarios, it is unclear, if both events happened simultaneously or if the data did not allow a precise distinction between the two. According to Ray and colleagues (2003) large spatial expansions can lead to the same signal in the mismatch distribution than a pure demographic expansion in a panmictic (= randomly mating) population. It is therefore likely that a single expansion happened in the past of the analysed populations. However, it was not possible to date the expansion (see below).
A study on mitochondrial D-loop of *M. murinus* in the same study site as our project reported two recent spatial expansions and dated them between 26,500 to 33,500 years ago and 3,300 to 14,000 years ago (Schneider et al. 2010). Interestingly, we did not see bimodal mismatch distributions in any species for any locus. However, the two studies are not easily comparable as they use genomic or mitochondrial DNA, respectively. The haplotypes of the D-loop of *M. murinus* in the study area showed pairwise differences of 25 bp to 29 bp in the second mode (the older mode) of the mismatch distribution (Schneider et al. 2010, "JBA" in their Fig. 2). Such high pairwise difference was not seen in *V1R* s of *M. murinus* where the highest nucleotide difference between haplotypes was 5 bp in *VN1R Mmur048*. This was also shown by the much lower nucleotide diversity in the *V1Rs* in *M. murinus* (0.0% – 0.19%, Table 5.2) compared to the D-loop (1.8%, Schneider et al. 2010, their Table 1) and the diversity of *V1Rs* seemed to be too low within the population to detect a bimodal distribution. However, it has to be mentioned that this effect might be biased by the primer design where – although rarely – highly diverse amplicons led to the rejection of the initially designed primers or to the rejection of the whole locus. The mismatch distributions showed higher variation in *M. ravelobensis* for several *V1R* genes. This was also shown by the high number of haplotypes, the higher haplotype and nucleotide diversity, and the higher number of polymorphic sites in *M. ravelobensis* compared to *M. murinus*. Demographic equilibrium has been reached at least in the three loci with ragged distributions. Our τ-values in VRs were smaller than in the D-loop but purifying selection could have led to an underestimation of τ (see below). The τ-values were higher in *M. ravelobensis* indicating that the potential expansion of *M. ravelobensis* is older than in *M. murinus*. However, because of the highly variable results for the different VR loci we did not calculate potential times of the expansions for the two species. In summary, *M. murinus* shows patterns for a recent decrease in genetic diversity by a founder effect or genetic bottleneck and a subsequent expansion. The last expansion in *M. ravelobensis* likely happened earlier and some loci actually show patterns of demographic equilibrium. The low diversity in several loci could be related to purifying selection as addressed below.

The most recent estimated time of an expansion in *M. murinus* (3,300 years ago, Schneider et al. 2010) indicated that this expansion did not follow an anthropogenically caused genetic bottleneck, since humans arrived on the island of Madagascar about 2,000 years ago (Green and Sussman 1990). *M. murinus* has a wide distribution from southern to northern Madagascar (Mittermeier et al. 2008). However, ecologically it prefers dry habitats typically
seen in southern Madagascar (Rakotondravony and Radespiel 2009), where also its most closely related sister species, *M. griseorufus*, is native (Yoder et al. 2000; Olivieri et al. 2007). It is therefore highly likely that *M. murinus* rather recently colonised the study area and the observed patterns of low genetic diversity might be explained by a founder effect.

**Purifying selection in the recent evolutionary history of vomeronasal receptor genes**

It was shown that the majority of *VIR* gene clusters in mouse lemurs evolved under strong positive selection and repeated gene duplication led to the evolution of a large *VIR* repertoire (Hohenbrink et al. 2012). Positive selection still acted on *VIR* genes during the diversification of mouse lemurs as indicated by analyses of single *VIR* loci across different mouse lemur species (Hohenbrink et al. 2012). The positive selection, which was probably involved in generating this high diversity, was expected to be ongoing. With sequence data from 20 individuals per population and species we were able to look at selection still acting on present-day populations. The results suggested that *VIR* genes may currently rather evolve under purifying selection. Purifying selection reduces variation in the population and secures the functional stability of the receptor. Several results supported the presence of purifying selection (or the lack of positive selection) on the population level: 1) McDonald-Kreitman tests were not significant (with one exception). The test calculates if polymorphic differences (= differences within the tested species) are based on an excess of nonsynonymous substitutions compared to the fixed differences (= differences between the two sister species). Positive selection was therefore not confirmed for the majority of loci. 2) Nonsynonymous substitutions were randomly distributed within the *VIR* protein indicating neutral evolution rather than positive selection. Positive selection would have targeted mutations on specific protein domains and especially the ligand binding site, but here the frequency of substitutions was not increased. 3) The most striking argument is the high number of *d*$_s$ compared to *d*$_N$. A *d*$_{S}$/*d*$_s$ ratio below 1 is strong evidence for purifying selection and was found in most loci of both species. However, purifying selection is probably not equally strong as the highly diverse results of the neutrality tests suggest variable evolutionary histories in different *VIR* loci.

In *M. ravelobensis* VN1R Mmur066 is a special case of the 15 selected *VIR* loci. It is the only locus with significant MKT and it has the largest number of haplotypes. These results are consistent with positive selection acting on this locus in this species, but possibility of a false positive result cannot be rejected. The largest genetic distance between two haplotypes
in VNIR Mmur066 was six substitutions. Like Mmur066, the $d_S/d_S$ ratios of Mmur033 and 040 in M. ravelobensis were above 1 and according to the mismatch distributions with high variability (App. 5.2) those two loci might have also evolved under positive selection. The only $d_S/d_S$ ratio above 1 in M. murinus (VNIR Mmur043) seemed to be a statistical artefact due to the low number of substitutions considering the half-bell shaped mismatch distributions. A comparison of genetic diversity and selection pressures between locus vs. corresponding cluster (see Hohenbrink et al. 2012) or between clustered vs. unclustered locus did not reveal any pattern.

**Conclusion**

The VR diversity of M. murinus and M. ravelobensis in northwestern Madagascar has been shaped by processes of population expansion and purifying selection. The signals of strong positive selection found in the whole VIR repertoire and its separate monophyletic gene clusters have led to a high diversification among mouse lemurs in distant history (Hohenbrink et al. 2012). However, more recently selection pressures may have shifted towards purifying selection in the majority of VIR loci to maintain the initial function of individual receptors. This study only analysed a small subset of the large VR repertoire but gives important insights into the recent evolution of VRs in these small nocturnal primates. The functional VR loci were not best-suited for demographic modelling considering the difficulty of differentiating between signals of purifying selection or recent genetic bottlenecks in several loci. However, the analysis of synonymous and nonsynonymous substitutions helped to understand the evolution of these loci and future studies should also consider functional loci.

We thank the Institute of Animal Breeding and Genetics at the University of Veterinary Medicine Hannover for their technical help. We are particular grateful to Sandra Thorén, Pia Eichmüller and Sarah Hohenbrink who collected the tissue samples in Madagascar. All field handling and sampling procedures adhered to the legal requirements of Madagascar and were approved by the Ministry of Water and Forests. We have complied with the ethical standards for the treatment of primates and with national laws and research rules formulated by the Malagasy authorities. This work was funded by the Volkswagen Foundation (grant number I/84 798).

5.5 References


5.6 Appendix

App. 5.1: IDs for the sampled individuals from Madagascar (JBA); F = female; M = male

<table>
<thead>
<tr>
<th>Sex</th>
<th>ID</th>
<th>Sex</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>02-05</td>
<td>F</td>
<td>30-06</td>
</tr>
<tr>
<td>M</td>
<td>32-06</td>
<td>M</td>
<td>12-07</td>
</tr>
<tr>
<td>F</td>
<td>05-07</td>
<td>M</td>
<td>30-07</td>
</tr>
<tr>
<td>F</td>
<td>20-07</td>
<td>M</td>
<td>31-07</td>
</tr>
<tr>
<td>M</td>
<td>23-07</td>
<td>F</td>
<td>51-07</td>
</tr>
<tr>
<td>M</td>
<td>57-07</td>
<td>F</td>
<td>68-07</td>
</tr>
<tr>
<td>F</td>
<td>58-07</td>
<td>F</td>
<td>74-07</td>
</tr>
<tr>
<td>F</td>
<td>60-07</td>
<td>M</td>
<td>02-08</td>
</tr>
<tr>
<td>M</td>
<td>75-07</td>
<td>M</td>
<td>10-08</td>
</tr>
<tr>
<td>F</td>
<td>01-08</td>
<td>F</td>
<td>14-08</td>
</tr>
<tr>
<td>M</td>
<td>08-08</td>
<td>M</td>
<td>41-08</td>
</tr>
<tr>
<td>F</td>
<td>13-08</td>
<td>F</td>
<td>42-08</td>
</tr>
<tr>
<td>M</td>
<td>17-08</td>
<td>M</td>
<td>44-08</td>
</tr>
<tr>
<td>M</td>
<td>24-08</td>
<td>F</td>
<td>49-08</td>
</tr>
<tr>
<td>F</td>
<td>28-08</td>
<td>F</td>
<td>51-08</td>
</tr>
<tr>
<td>M</td>
<td>34-08</td>
<td>M</td>
<td>52-08</td>
</tr>
<tr>
<td>F</td>
<td>35-08</td>
<td>F</td>
<td>68-08</td>
</tr>
<tr>
<td>M</td>
<td>57-08</td>
<td>M</td>
<td>69-08</td>
</tr>
<tr>
<td>M</td>
<td>64-08</td>
<td>F</td>
<td>70-08</td>
</tr>
<tr>
<td>F</td>
<td>67-08</td>
<td>F</td>
<td>72-08</td>
</tr>
</tbody>
</table>
App. 5.2: Observed and simulated mismatch distributions of *M. murinus* (Mmur, left side, grey) and *M. ravelobensis* (Mrav, right side, orange); simulated (d) = simulated after demographic expansion model (line with circles), simulated (s) = simulated after spatial expansion model (line with crosses); (continues on next pages)
App. 5.2: continued
App. 5.2: continued
App. 5.2: continued
6. General discussion

The evolution of the VR repertoire in mouse lemurs was analysed in four separate studies investigating the present composition of the \( V1R \) and \( V2R \) repertoire, their expression in the two olfactory organs and their intrapopulation and interspecific variation. This general discussion tries to connect the four studies to get a bigger picture of the evolution of VR-mediated olfactory communication in mouse lemurs.

6.1 Clusters and families

In chapter 2 nine monophyletic clusters were identified in the \( V1R \) repertoire of the grey mouse lemur together with ten unclustered loci (Hohenbrink et al. 2012). Per definition the loci of one cluster had to have evolved from a single locus in the ancestral lemur and diversified and expanded later during lemur or mouse lemur evolution. \( V1R \) sequences of Euarchontoglires species were added to the sequences of the grey mouse lemur – including other primates and their sister taxa (tree shrews, lagomorphs and rodents) – to analyse if true monophyly is given. The data were derived from Young et al. (2010). In that study, 1809 intact \( V1R \)s were found across 37 mammalian species. Using only Euarchontoglires species reduced the sample to 1049 sequences (992 after removing incomplete data) and allowed a more precise alignment of the \( V1R \) data. A phylogenetic reconstruction in chapter 2 revealed that subsequently defined clusters are separated by heterospecific sequences indicating a split of the ancestral sequences of two clusters before the differentiation of mouse lemurs or lemurs in general. The grey mouse lemur was the only lemur represented in the analysis and it seemed appropriate to consider that some or all clusters are specific for lemurs. Therefore, the \( V1R \) cluster classification should not be affected by newly published \( V1R \) sequences in lemurs: Yoder and colleagues (2014) recently published several new \( V1R \) sequences of cluster I among a variety of lemur species. The monophyly was not supported anymore for mouse lemurs but was still true for the common lemur ancestor.

Before comparing \( V1R \) clusters of grey mouse lemurs with \( V1R \) families in rodents, a comparison of the classification method for the two terms is needed. Mouse \( V1R \) families and later \( V2R \) families as well have not been defined by monophyly (which would have required the VR data of several non-murine mammals) but by the similarity of amino acid sequences (Rodriguez et al. 2002; Yang et al. 2005). Two guidelines have been established: 1) Members of a VR family share at least 40% of their amino acid sequences. 2) Between VR families the similarity of amino acid sequences is below 40%. Whereas the first guideline is strict and true
for all 12 V1R families in mice, the second guideline does not fit for most families since 10 of 12 families show interfamilial similarities slightly above 40% with up to 55% similarity between the families V1Ra and V1Rb (Rodriguez et al. 2002, their Fig. 2). However, according to the authors these families are separated because fusions of two of them would have led to families with intrafamilial similarities below 40%. All V1R genes in mice share at least 15% of their amino acid sequences.

In order to improve comparability between mouse lemur and rodent datasets, the amino acid sequence similarities within and between clusters of mouse lemurs were calculated to classify them into families as in rodents (data not shown before). Cluster IV and VNIR Mmur066, cluster VI and VII as well as cluster VIII and IX show more than 40% similarity, respectively, and would form families according to the guidelines in rodents. A fourth family would be cluster V that is not closely related to any other cluster. Nevertheless, following these guidelines it would not be possible to group clusters I, II and III (plus the remaining unclustered loci) into one family. In some cases the pairwise differences in amino acid sequence similarity between two loci was as low as 35.9% which would not allow grouping them into a single family. Additionally, the group could not be divided without splitting closely related loci. However, all those clusters that were defined in chapter 2 are distinct in the Neighbour Joining tree (Fig. 2.1). When looking at similarities within and between V1R clusters (Fig. 6.1), there is no overlap of amino acid sequence similarity within clusters and between single clusters and the remaining repertoire. The similarity within clusters ranged from 66.4% to 99.3%, whereas the highest similarity to another locus of the repertoire was 62.0% between cluster III and the unclustered VNIR Mmur076. Without unclustered loci the similarities between clusters are even below 60%. All loci of the whole V1R repertoire share at least 20% of their amino acid sequence. The maximum diversity is therefore slightly lower in mouse lemurs than in mice.
Fig. 6.1: Range of similarities of the amino acid sequences (in percent) within V1R clusters and between the sequences of single V1R clusters and all remaining V1R sequences (including unclustered loci); only 105 V1R loci from Young et al. (2010) were included as described in chapter 2.

With the exception of cluster I, which was distinct in mouse lemurs, all other clusters have sister clusters/families in mice. Clusters II and III are embedded within the mouse V1Ref family, whereas clusters VIII and IX are both most closely related to family V1Rc. The clusters IV, V, VI, and VII are most closely related to mouse families V1Rd, V1Rjk, V1Ri, and V1Rh, respectively. The division of the V1R repertoire into clusters as done in chapter 2 seems therefore to be appropriate considering the distribution of the similarities of the amino acid sequences without overlap within and between clusters (see Fig. 6.1). Together with the organisation of V1R loci in analogy to mice, it is likely that the clusters can be considered as “families” despite the 40% similarity cut-off guideline known from mice. The cut-off for the grey mouse lemur would need to be at 65% similarity. Two methods to classify gene clusters have been shown in this section: The advantage of using amino acid sequence similarities is the independent use without V1R data from other species. The advantage of monophyly is that such gene clusters have evolved from a single gene within a given taxon. Such evolutionary unit is suited for selection modelling shown in chapter 2. Considering the growing availability of VR repertoires, future classifications should consider both classification methods.
6.2 The repertoires of vomeronasal receptors in retrospective

In this thesis the term “locus” was used for all V1R and V2R sequences equivalently to the term “gene” and all studies of this thesis handled the VR sequences as if they were separate genes. However, it was often not possible to confidently distinguish between alleles and genes. Highly similar sequences could have been separate genes after recent gene duplication or alleles of the same gene. More diverse sequences could have been two genes or two highly distinct alleles of the same gene. It was therefore necessary to know if two sequences with a set number of nucleotide differences were two separate genes or two alleles of the same gene. This problem and the potential number of intact VR genes in the mouse lemurs are discussed in this section.

Young and colleagues (2010) published 107 intact V1R sequences for the grey mouse lemur and predicted a total of 214 intact V1R genes to exist in the whole genome. This prediction was based on their estimation that the existing dataset contained only about one half of the total genome sequence information, and it was assumed that the other half should contain an equal amount of V1R sequences. However, there are two problems with this argument: 1) the initial number of V1R sequences might already include multiple alleles of the same genes, and 2) the estimation might not be correct because of an unequal distribution of V1Rs between the published and the unpublished half of the genome. Whereas the first problem leads to an overestimation of the number of genes, the second problem could cause and over- or underestimation of the number of genes. It can be asked whether a better estimation of the repertoire may be possible after having completed this thesis. Two of the published sequences had already been excluded in all studies of this thesis because of identical amino acid sequences (VN1R Mmur102) or incomplete coding sequence length (VN1R Mmur073), but further loci were probably wrongfully assigned to independent genes. As a starting point, a threshold is needed to separate genes and alleles. Young and colleagues (2010) already used a conservative subset where all sequences with a nucleotide identity ≥98% where removed except for the longest sequence. The conservative subset consisted of 87 instead of 107 V1R genes. Depending on the total length of the V1R sequence, 98% nucleotide identity means that a difference between two sequences above 18-20 bp is needed to be counted as two genes. Chapter 5 showed that the maximum nucleotide distance between two haplotypes (∼ alleles) of the same locus was 16 bp (seen in VN1R Mmur040 in M. ravelobensis). Therefore, the conservative subset with 87 genes seems to be able to exclude even high allelic variation. High allelic variation was especially found in the golden-brown...
mouse lemur, *M. ravelobensis*, but similarly high variation is expected for natural populations of the grey mouse lemur that did not undergo a recent genetic bottleneck or founder effect as indicated in the analysed population of the Ankarafantsika National Park (chapter 5). For example, study 2 showed that the \( \frac{d_{S}}{d_{S}} \) ratios in the grey mouse lemur lineage (represented by one northern and one southern *M. murinus* individual) were higher than in the red lineage for all seven tested individual *V1R* genes indicating that high variation is also possible for the grey mouse lemur. In chapter 4 it was reported that 64 different *V1R* loci were expressed and all of them could be individually distinguished by using locus-specific primer pairs. The minimum nucleotide distance between the expressed loci was 28 bp (= about 97% nucleotide identity), therefore all expressed loci can be considered as true genes. A minimum number of 64 *V1R* genes in the genome of the grey mouse lemur can therefore be assumed as a conservative estimate. After the design of locus-specific primer pairs in chapter 4 some primers did not amplify the targeted loci but revealed new and previously unpublished loci (data not shown). The primer pairs were discarded, but this finding indicates that unknown loci are present in the genome of mouse lemurs, and can probably be revealed with the help of full genomic data. Once the full genome has been published for the grey mouse lemur, a threshold of 98% nucleotide identity seems to be appropriate to distinguish genes from alleles. Based on these considerations, the predicted number of 214 intact *V1R* loci is most likely an overestimate (because of the inclusion of some alleles rather than genes). The true repertoire of *V1R* genes probably lies rather between 70 and 200 intact genes in the grey mouse lemur.

Among mammals mouse lemurs possess one of the largest *V1R* repertoires. They are therefore potentially able to perceive a large number of different ligands. The vomeronasal system mainly detects pheromones used for intraspecific communication or kairomones used for predator recognition and avoidance (see Isogai et al. 2011). Both classes of semiochemicals are highly relevant for mouse lemurs. Mouse lemurs show a variety of different marking behaviours (Glatston 1979), for example, to mark their sleeping sites (Braune et al. 2005) and perceive the reproductive status of females by olfaction (Buesching et al. 1998). On the other hand, mouse lemurs suffer from a high predation risk (Scheumann et al. 2007) and potentially recognise terrestrial mammals by olfactory cues (Sündermann et al. 2008; Kappel et al. 2011). An enlargement of the *V1R* repertoire by diversification can therefore be assumed to be highly beneficial for mouse lemurs, as it allows complex olfactory perception and communication. Wang and colleagues (2010) showed that nocturnal mammals possess significantly more functional *V1Rs* than diurnal mammals. Olfactory signals allow
intraspecific communication in nocturnal mouse lemurs even when vision is limited during darkness. However, nocturnal activity alone cannot explain the large number of V1R genes in mouse lemurs, because the galago’s repertoire was estimated to contain less than half as many genes (Young et al. 2010) despite the similar body size and the nocturnal and arboreal activity (Bearder 1987).

During evolution the V1R repertoire has increased in size after several gene duplication events. Gene duplications most likely increase the number of genes on the same chromosome and in mice many large V1R families (including the largest families V1Rc and V1Rd) are located on the same chromosome (Zhang et al. 2004; Young et al. 2010). The majority of V1Rs is found on the chromosomes 6, 7, 13 and 17 but there are also single genes on some chromosomes (in total eight chromosomes contain intact V1Rs). In mouse lemurs a similar pattern is expected where genes of the same cluster are potentially located on the same chromosome.

It is likely that some unclustered genes form undetected monophyletic gene clusters with yet unknown V1R genes that could potentially be revealed after the publication of the complete genome for the grey mouse lemur. The best candidate for a new cluster would be VN1R Mmur066 which was under significant positive selection in the analysed population of M. ravelobensis and showed the highest number of haplotypes (chapter 5). Most importantly a sister cluster of this gene is present in O. garnettii which consisted of 10 loci and has also evolved under significant positive selection (App. 2.5 A). Similar gene duplications and diversifications are therefore likely in mouse lemurs or other lemurs. However, the comparison of the V1R repertoires of M. murinus and O. garnettii in chapter 2 revealed that M. murinus has higher numbers of V1Rs per cluster rather than an increase in the number of clusters indicating a certain level of conservation in the number of clusters in strepsirrhines.

V2Rs had been reported to be completely pseudogenised in primates, although information outside catarrhine primates was missing (Young and Trask 2007). However, two V2Rs were shown to be expressed in the grey mouse lemur (and were present in the genomes of two other strepsirrhines, see chapter 3). Following the rational of Young and colleagues (2010) for mammals with incomplete genomes, about 4 estimated intact loci can be expected to be present in the genome of mouse lemurs. Because of a nucleotide identity of only 47%, the two expressed loci were considered as separate genes that even can be regarded as representatives of two different families (amino acid sequence similarity of 33%).
genes were most closely related to the mouse V2R families C or D, respectively, which were the two most basal V2R families in mice (Fig. 3.1). The revised estimated VR repertoires in primates after the study in chapter 3 are shown in Fig. 6.2.

![Fig. 6.2: Phylogenetic reconstruction of selected primates and estimation of functional V1R and V2R genes in the genomes of primates with data of this thesis shown in red (based on Young and Trask 2007; Young et al. 2010, Fig. 2; and Hohenbrink et al. 2013); ? = no information available]

Whereas mouse lemurs have large repertoires of V1Rs, so far only two V2Rs have been found (Young et al. 2010; Hohenbrink et al. 2013). This finding indicates a high discrepancy between the VR classes. Rodents also have larger V1R than V2R repertoires but the level of difference is not comparable to mouse lemurs. Mice have 239 V1Rs and 123 V2Rs, whereas rats have 108 V1Rs and 87 V2Rs according to estimates (Young and Trask 2007; Young et al. 2010). The low number of V2Rs in all three analysed strepsirrhine primates indicated that 1) V2Rs did not differentiate much in strepsirrhines, or 2) the majority of V2Rs have still not been detected in the genome. The genomes had a coverage of 2x for M. murinus (Genbank ABDC 0000 0000.1), a coverage of 2x for O. garnettii (AAQR 0000 0000.3), and a coverage of ~38x for the aye-aye, D. madagascariensis, (AGTM 0000 0000.1). However, even the high coverage genome of D. madagascariensis did not reveal additional intact loci. Although divergent evolutionary histories of the aye-aye and mouse lemurs are likely, no large V2R repertoires would be expected in mouse lemurs based on the data of the closely related aye-aye. During BLAST searches in the genomes of strepsirrhines only one V2R pseudogene was found in D. madagascariensis and two V2R pseudogenes were present in the O. garnettii genome (data not shown before). Interestingly, no V2R pseudogenes were found in M. murinus, but notably, 153 V1R pseudogenes were reported (Young et al. 2010), a
larger number than the number of intact \textit{V1R} loci. The low number of pseudogenes indicated that the number of \textit{V2R} loci has probably never increased much during strepsirrhine evolution. \textit{V2R}s in rodents bind large peptides like MUPs (Chamero et al. 2007) or MHC (Leinders-Zufall et al. 2004). Female mouse lemurs are assumed to exercise post-copulatory mate choice based on MHC constitution (Schwensow et al. 2008), and MHC detection is potentially mediated by \textit{V2R}s similar to rodents. However, further research will be necessary to identify the function of \textit{V2R}s in mouse lemurs.

6.3 Selection pressures on vomeronasal receptors: a revision

Two different modes of selection are likely for VRs: Purifying or positive selection. If the function of a receptor has to be conserved, purifying selection will be the dominant evolutionary force. In receptors with well adapted function nonsynonymous substitutions, which would change the function of the receptor, would be disadvantageous. Selection against nonsynonymous substitutions leads to an excess of synonymous substitutions ($d_N/d_S < 1$). In contrast, a change in the function of the receptor might be advantageous in certain cases, and would cause positive selection. Here, nonsynonymous substitutions are favoured and accumulated ($d_N/d_S > 1$). For example, after gene duplication one gene could rapidly evolve to serve a different function or support the function of the previous gene. A random distribution of synonymous and nonsynonymous substitutions ($d_N/d_S = 1$) would indicate neutral evolution that is present in pseudogenes and not expected for the analysed functional genes.

The \textit{V1R} repertoire in mouse lemurs has evolved under various selection pressures. The early evolution of lemurs started approximately 62-65 million years ago (Yoder and Yang 2004) when they arrived on the island of Madagascar and had separated from lorisoids. The evolution of mouse lemurs started approximately 15-26 million years ago after the split from the most closely related extant genus \textit{Mirza} (Yoder and Yang 2004). Strong positive selection on the \textit{V1R} sequences was detected in the clusters of the grey mouse lemur in chapter 2. However, with the present dataset it was not possible to distinguish if the detected positive selection in \textit{V1R}s happened more prominently during the beginning of lemur evolution, after their radiation on the island or after the most recent split of the two genera \textit{Microcebus} and \textit{Mirza}. Therefore, no predictions can be made, if the extents of positive selection (together with a divergent \textit{V1R} repertoire) are generally found in lemurs, the family of cheirogaleids (which includes mouse lemurs) or specifically in mouse lemurs. Regardless, positive selection was still the prominent mode of selection during the radiation of the genus
Microcebus (8-12 million years ago, Yoder and Yang 2004) as indicated by the analyses of single genes (chapter 2). Only recently most genes evolved under purifying selection which was suggested in several individual loci (chapter 5). More precisely, purifying selection had always acted on the majority of sites since the overall $\omega$ M0 was below 1 in all clusters (Table 2.1) and individually tested loci (Table 2.2). However, the data of single populations showed that more recently (during the evolution within species) synonymous substitutions were accumulated and also the mismatch distributions (chapter 5) showed low variation in 11 of 15 V1R loci in M. murinus and 6 of 15 V1R loci in M. ravelobensis. It can be suggested that the mode of selection potentially varied depending on their function. In chapter 2 it was shown, that cluster IV without significant evidence for positive selection most likely binds female cues (the function was indicated in the mouse sister family by Isogai et al. 2011), which are probably strongly conserved. In contrast, cluster VI did show significant evidence for positive selection and binds male cues that may evolve rapidly by sexual selection (either by male-male competition or female choice). There was no evidence that the strength of selection in a particular cluster was correlated with the size of the cluster (chapter 2).

The position of the ligand binding site of the V1R protein is debated (see Hohenbrink et al. 2012 or chapter 2 of this thesis; Yoder et al. 2014). Protein models with confirmed binding sites are not available for V1Rs. The first of the two recent approaches to estimate the position of the binding site uses the distribution of sites under significant positive selection (chapter 2). The problem was that only 11 codons were found to be under significant positive selection throughout the whole V1R repertoire and, considering that 9 of them were located in the 2\textsuperscript{nd} or 3\textsuperscript{rd} extracellular loop, it was assumed that the ligand binding site is potentially generated by these two loops or the extracellular domains in general. However, Yoder and colleagues (2014) argued that the ligand binding site should include the 4\textsuperscript{th} and 5\textsuperscript{th} transmembrane region and the in-between 2\textsuperscript{nd} extracellular loop. Their suggestion was based on protein structure predictions. Both studies agreed in the inclusion of the 2\textsuperscript{nd} extracellular loop and because 7 of the 9 extracellular codons under significant positive selection were located in this loop, the presented results in chapter 2 are potentially valid in view of this new hypothesis. Because of the short N-terminal domain in V1Rs, it was previously hypothesised that the ligand binding site lies within the seven transmembrane domains (Kristiansen 2004), but evidence is still lacking. In human and mouse ORs, Man and colleagues (2009) found 22 sites that were conserved among putative orthologues but variable among paralogues. The sites were located in the transmembrane regions 2 to 7 and the 2\textsuperscript{nd} extracellular loop. An OR
model showed that these sites mostly face the inner lumen of a structural pocket that was suggested to be the ligand binding site. Another approach in ORs used point-mutated variants of a single receptor to identify amino acids essential for ligand binding (3rd, 5th and 6th transmembrane region involved in their example: Katada et al. 2005). In conclusion, the exact location of the ligand binding site is still unknown due to a lack of structural information on the receptor proteins (such as a high resolution crystal structure) and more research is needed to identify ligand binding sites of V1Rs that even may vary between distantly related V1Rs.

The V2R proteins are structurally similar to the metabotropic glutamate receptors mGluRs – including the taste receptors type 1, T1Rs – and it was therefore suggested that the long extracellular N-terminal domain is used for ligand binding (Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). Before, Takahashi and colleagues showed (1993) that chimeric mGluR1 receptors respond like regular mGluR2 receptors when at least the first half or more of the N-terminus of mGluR1 is replaced by the corresponding amino acid sequence of mGluR2. The study indicated that the first ~355 amino acids of the N-terminal are responsible for the ligand binding. Transferring the results to mouse lemurs the position of the ligand binding site in the two V2R genes could be predicted (data not shown before). In VN2R1 the first 355 amino acids of the external N-terminal domain took ~39% of the total amino acid sequence. It was found that ~23% (M. murinus) or ~32% (M. ravelobensis) of the total nonsynonymous substitutions within the respective species were located in this potential ligand binding site. Also the nonsynonymous substitutions between the two species (~23%) were not accumulated in this region indicating no evidence for positive selection targeting the ligand binding site of this V2R. For VN2R2 ~38% (M. murinus) or ~50% (M. ravelobensis) of the total nonsynonymous substitutions were located in the potential ligand binding site, which does not indicate overrepresentation (43% predicted). However, nonsynonymous substitutions between the two species were significantly accumulated in the potential ligand binding site (observed vs. expected: $\chi^2 = 4.6$, df = 1, $p = 0.032$). Nevertheless, similar to V1Rs structural information about V2Rs is lacking and further research is needed to interpret the distribution of nonsynonymous sites.

6.4 Overall conclusion

At least 18 different V1R genes were potentially present in the ancestral lemur and 9 of them diversified and expanded to gene clusters within lemurs. Given the situation in the strepsirrhine galago one further gene cluster could be expected within lemurs. Based on
expression data the V1R repertoire has to consist of a minimum of 70 genes and up to 200 are likely. A significant proportion of sites in the V1R repertoire evolved under positive selection. However, the finding of purifying selection as the prominent mode of selection during the most recent evolution of V1Rs indicates a potential decline in the diversification of V1Rs with all relevant functions being covered by the present repertoire or simply indicates that not all V1Rs are under positive selection all of the time. On the other hand, based on data from a closely related lemur species and the small published V2R repertoire, V2Rs did not diversify in mouse lemurs to a similar extent. The potential V2R repertoire probably contains less than 10 genes. V1Rs and V2Rs are expressed in the sensory organs of both olfactory systems in the grey mouse lemur. Given the differences in the neural pathways of MOE and VNO signals, which project to higher cortical brain centres or the limbic system, respectively, this raises the intriguing possibility that the evolution of MOE-expression of VRs enabled mouse lemurs to adaptively diversify the processing of VR-encoded olfactory information.

It can be concluded, that mouse lemurs evolved a very complex vomeronasal system to adequately perceive and process the information of the olfactory signals from conspecifics or heterospecifics. However, information about specific corresponding ligands is currently lacking in mouse lemurs, which would be needed to combine the genetic results with the actual ecological function of VRs. Furthermore, the genome of the grey mouse lemur is incomplete and a high-coverage genome is necessary to identify the full repertoire of V1Rs and V2Rs. Additional genome data from other lemurs and lorisoids would help to identify VR repertoires among strepsirrhine primates and reveal if the complexity in mouse lemurs is unique or typical for lemurs.

6.5 References


Statement (Erklärung)


Die Kapitel 2, 3 und 4 wurden dem Formatstil dieser Dissertation angepasst. Dabei wurde auch das Format für die Angabe von Literatur vereinheitlicht und die Nummerierung der Abbildungen und Tabellen leicht verändert. Die Kapitel sind hier angegeben wie publiziert oder zur Publikation eingereicht mit den folgenden Änderungen:

In Kapitel 2 „Pervasive and ongoing positive selection in the vomeronasal-1 receptor (V1R) repertoire of mouse lemurs“ wurde das Supplementary Material in Appendix (oder App.) umbenannt. Die Anhänge Supplementary Material S3 und S4 wurden wegen ihrer Länge nicht mit abgedruckt, können aber über mbe.oxfordjournals.org/content/29/12/3807/suppl/DC1 eingesehen werden. App. 2.5 A wurde mit Vermerk korrigiert.

In Kapitel 3 „First evidence for functional vomeronasal 2 receptor genes in primates“ wurde das Electronic Supplementary Material in Appendix (oder App.) umbenannt. App. 3.3 wurde mit Vermerk korrigiert.

Kapitel 4 „Functional promiscuity in a mammalian chemosensory system: Extensive expression of vomeronasal receptors in the main olfactory epithelium of mouse lemurs“ wurde nicht weiter als oben angegeben verändert.


Institut für Zoologie, Tierärztliche Hochschule Hannover, Deutschland
Department of Zoology, University of Cambridge, Vereinigtes Königreich

Die Dissertation wurde bisher nicht für eine Prüfung oder Promotion oder für einen ähnlichen Zweck zur Beurteilung eingereicht. Ich versichere, dass ich die vorstehenden Angaben nach bestem Wissen vollständig und der Wahrheit entsprechend gemacht habe.

Datum, eigenhändige Unterschrift (Philipp Hohenbrink)
Acknowledgements (Danksagung)

Ich bedanke mich bei allen, die über 100 Seiten inklusive Anhang durchgelesen haben und nun endlich auf der letzten Seite angekommen sind, sowie all jenen, die nur mal eben nachschauen wollten, ob ihnen auch angemessen gedankt wurde, und gleich die letzte Seite aufgeschlagen haben.

Ganz besonderer Dank geht an Dr. Nick Mundy und Prof. Dr. Ute Radespiel (in alphabetischer Reihenfolge) für die Bereitstellung des Themas und die Begutachtung der Arbeit sowie die tolle Betreuung über die letzten Jahre. Die Kommentare und Diskussionen während dieser Zeit haben mir sehr geholfen. ("I especially would like to thank Dr Nick Mundy and Prof Dr Ute Radespiel (in alphabetical order) for the initial design and evaluation of this thesis as well as the excellent supervision during the last years. The comments and discussions during that time were a great help")

Bei Prof. Dr. Miguel Vences bedanke ich mich für die zweite Begutachtung dieser Arbeit.


Ich danke Steve, Neil und Oscar für die Zeit in Cambridge im kleinsten Büro der Welt. Vielen Dank an Ewa-Lilian und June für die Unterbringung und liebevolle Versorgung in Cambridge. ("I would like to thank Steve, Neil and Oscar for the time in our tiny whale-watching office in Cambridge. Special thanks go to Ewa-Lilian and June for their accommodation and lovely support in Cambridge")

Ich danke außerdem Dr. Henrike Hartmann und Dr. Oliver Grewe von der Volkswagen-Stiftung für die freundliche und kompetente Beantwortung aller Fragen sowie der Bearbeitung diverser Anträge. Der Volkswagen-Stiftung danke ich für die Finanzierung der kompletten Doktorarbeit.

Ich bedanke mich bei Janet M. Young, die 50x in dieser Arbeit zitiert wurde, und die besonders durch ihre Publikation des VIR Repertoires von Mausmakis viel für diese These getan hat. Dank geht an Anne D. Yoder für aktuelle Informationen über das langerwartete Mausmaki-Genom sowie E. Barry Keverne für das Kommentieren des vierten Kapitels.