




Genomes of *Microtus* Rodents Highlight the Importance of Olfactory and Immune Systems in Their Fast Radiation

Alexandre Gouy ^{1,2,†}, Xuejing Wang ^{1,2,†}, Adamandia Kapopoulou ^{1,2},
Samuel Neuenschwander ³, Emanuel Schmid ³, Laurent Excoffier ^{1,2,*}, Gerald Heckel ^{1,2,*}

¹Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

²Swiss Institute of Bioinformatics, Lausanne, Switzerland

³Vital-IT, Swiss Institute of Bioinformatics, Lausanne, Switzerland

[†]Equal contribution.

*Corresponding authors: E-mails: laurent.excoffier@unibe.ch; gerald.heckel@unibe.ch.

Accepted: October 07, 2024

Abstract

The characterization of genes and biological functions underlying functional diversification and the formation of species is a major goal of evolutionary biology. In this study, we investigated the fast radiation of *Microtus* voles, one of the most speciose group of mammals, which shows strong genetic divergence despite few readily observable morphological differences. We produced an annotated reference genome for the common vole, *Microtus arvalis*, and resequenced the genomes of 10 different species and evolutionary lineages spanning the *Microtus* speciation continuum. Our full-genome sequences illustrate the recent and fast diversification of this group, and we identified genes in highly divergent genomic windows that have likely particular roles in their radiation. We found three biological functions enriched for highly divergent genes in most *Microtus* species and lineages: olfaction, immunity and metabolism. In particular, olfaction-related genes (mostly olfactory receptors and vomeronasal receptors) are fast evolving in all *Microtus* species indicating the exceptional importance of the olfactory system in the evolution of these rodents. Of note is e.g. the shared signature among vole species on Olfr1019 which has been associated with fear responses against predator odors in rodents. Our analyses provide a genome-wide basis for the further characterization of the ecological factors and processes of natural and sexual selection that have contributed to the fast radiation of *Microtus* voles.

Key words: arvicolinae, voles, reference genome, genome scan, rapid evolution, rodent diversification.

Significance

The explosive radiation of *Microtus* voles has resulted in probably more than 65 species within less than 2 million years that display remarkably little morphological differences. The sequencing of a high-quality reference genome of the European common vole *M. arvalis* and resequencing of 10 different species and evolutionary lineages showed that genes related to olfaction, immunity and metabolism evolved particularly fast in this radiation. These analyses provide a genome-wide basis for the further characterization of the ecological factors and processes of natural and sexual selection that have contributed to the particularly rapid diversification of *Microtus* voles.

© The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Introduction

Understanding the processes leading to the formation of new species remains a major challenge in evolutionary biology. The speciation rates among different lineages are highly imbalanced, with certain lineages gaining exceptional diversity through species radiation (Losos 2010; Rabosky et al. 2013). Many factors contribute to the speed of diversification, including sexual selection (Coyne and Orr 2004; Wagner et al. 2012), ecological opportunities (Wagner et al. 2012), and the capacity to adapt to such opportunities (Rabosky et al. 2013). The theory of speciation predicts the evolution of pre and postzygotic reproductive barriers between species (Dobzhansky 1937; Mayr 1942). These barriers can evolve through genetic drift, when isolated taxa accumulate differences in traits contributing to reproductive isolation, but natural selection also plays a role in the formation of new species, e.g. by adaptation to different environments (Hatfield and Schluter 1999; Rundle and Nosil 2005; Nosil et al. 2009; Comeault et al. 2015; Meier et al. 2018; Marques et al. 2019) or by selection against hybrids reducing cross-fertilization between taxa (Haldane 1922; Naisbit et al. 2001; Price and Bouvier 2002; Coyne and Orr 2004; Turelli and Moyle 2007; Svedin et al. 2008; Stelkens et al. 2010). Major functions involved in both pre and postzygotic isolation have been determined. Sensory systems seem to play a key role both through ecological adaptation and sexual selection, e.g. in cichlid fishes (Stelkens et al. 2010), fruit flies (Grillet et al. 2012; Yukilevich et al. 2016), stick insects (Nosil 2007), or the house mouse (Smadja et al. 2004; Hurst et al. 2017).

Genome sequencing of different species recently led to a better understanding of the genetic architecture and biological functions underlying the speciation process (Martin et al. 2013; Wolf and Ellegren 2017; Meier et al. 2018; Marques et al. 2019). Genome scans to detect highly divergent regions have identified many candidate genes potentially involved in reproductive isolation (Ellegren et al. 2012; Poelstra et al. 2014; Wolf and Ellegren 2017). However, genomic divergence is not just shaped by reproductive isolation, but also depends on intrinsic features such as recombination and mutation rates, or the density of variants under purifying selection (Cruickshank and Hahn 2014; Burri 2017). It remains a challenge to control the false-positive rate for genome scans (Teshima et al. 2006; Thornton and Jensen 2007; Soni et al. 2023). It is therefore important to consider the interplay between these forces when interpreting landscapes of genomic divergence (Wang and Heckel 2024).

Rodents of the *Cricetidae* family are among the most species-rich mammals. The genus *Microtus* is particularly interesting for studying diversification because its explosive radiation generated at least 65 species across the world in less than 2 million years (Jaarola et al. 2004; Fink et al.

2010; Beysard and Heckel 2014; Lischer et al. 2014). Indeed, despite their morphological similarity, they present much higher levels of genetic differentiation than other fast-radiating vertebrates, suggesting ongoing speciation processes and the presence of cryptic species (Heckel et al. 2005; Bastos-Silveira et al. 2012; Beysard et al. 2012; Paupério et al. 2012; Sutter et al. 2013). Furthermore, complete or partial reproductive isolation has been observed both between (De Jonge and Ketel 1981; Pierce et al. 1989; Soares 2013; Torgasheva and Borodin 2016) and within species, e.g. in the common vole *Microtus arvalis* (Heckel et al. 2005; Beysard and Heckel 2014; Beysard et al. 2015). It has been suggested that mechanisms such as sexual selection based on olfactory signals (Soares 2013; Duarte et al. 2016; Cerveira et al. 2019) played a central role in vole species divergence, but the genomic bases of these speciation processes remain unclear as no full-genome study has been performed so far in *Microtus* rodents.

In this study, we sequence and assemble *de novo* a high-quality, chromosome-scale, reference genome of the common vole, *M. arvalis*, which is used to resequence 12 individuals from 10 vole species and evolutionary lineages spanning the divergence continuum of the genus and an outgroup, *Myodes*. We characterize patterns of diversity and divergence along these genomes and use a new approach to identify regions of high divergence between species, as well as fast-evolving genes which are likely of particular relevance for the radiation of *Microtus* voles. Our results indicate an exceptional role of olfaction in the fast divergence of *Microtus* rodents, together with immunity and metabolic processes.

Results

A Chromosome-level Reference Genome for *Microtus arvalis*

We generated a reference genome of common vole *M. arvalis* (MicArv1.0) using Chicago and Hi-C scaffolding approaches (see Materials and Methods). We obtained a chromosome-scale assembly with 5 scaffolds larger than 150 megabases (Mb), and 17 smaller ones, shorter than 100 Mb (Fig. 1, supplementary fig. S1, Supplementary Material online), which is in very good agreement with the species' known karyotype showing 5 large and 17 smaller autosomes (Gamperl 1982). Note that X and Y sex chromosomes are both absent from the assembly of this male individual because the coverage was too low compared to the rest of the genome to be included in the initial scaffolding and assembly. Hereafter, we will therefore focus on the 22 largest scaffolds corresponding to the 22 autosomes and we will number them according to their sizes from 1 (largest) to 22 (smallest).

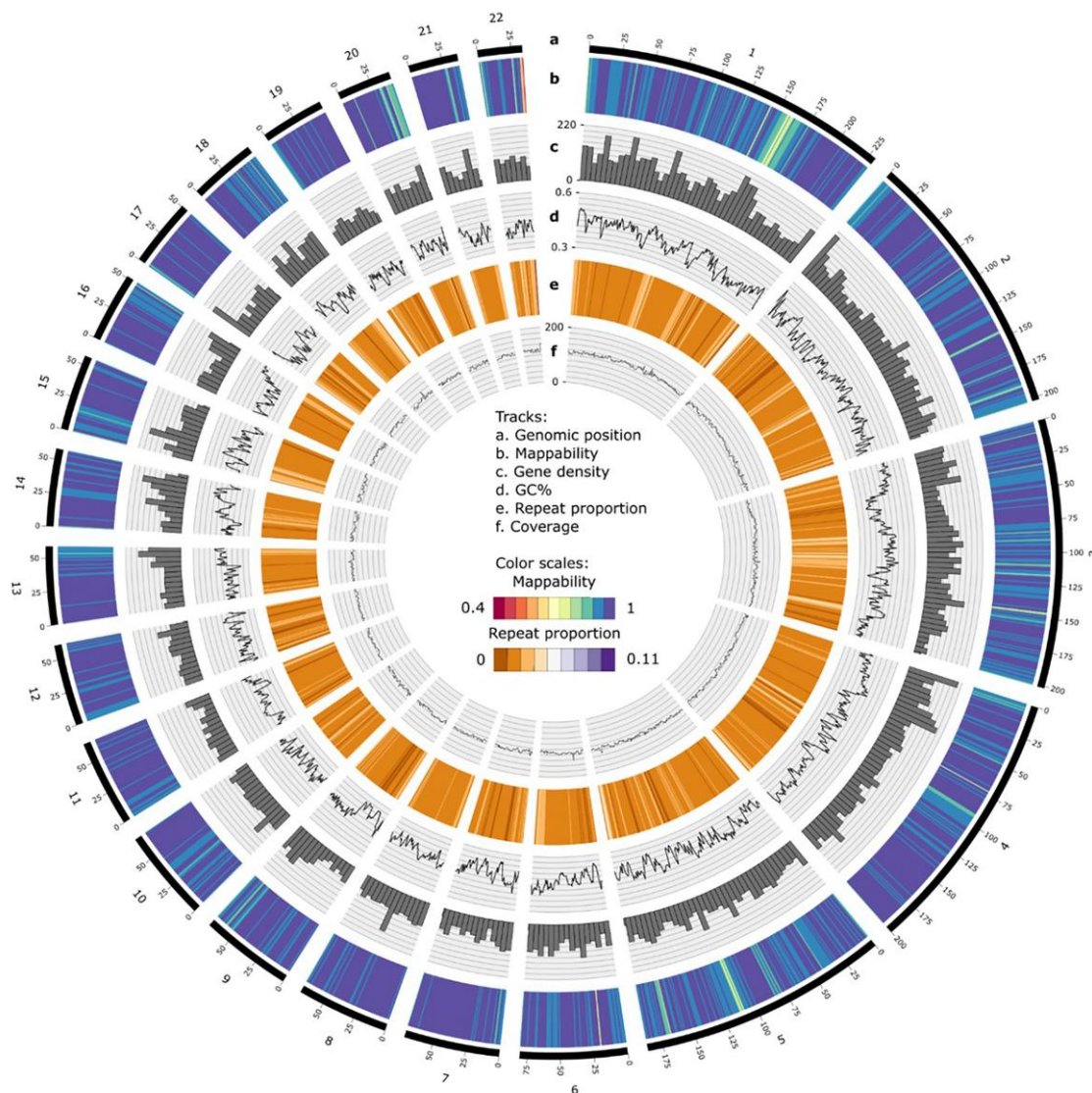


Fig. 1. Circos plot representing basic metrics along the *M. arvalis* genome computed over 1 Mb windows (except for gene density that is represented over 5 Mb windows). The outer track a) represents genomic position in Mb along the 22 largest scaffolds. Then, from the most external to the most internal track, we report: the mappability b), the number of genes per 5 Mb bins c), GC content d), proportion of repeats e), and depth of coverage of the sequenced individual f).

We compared the genome of *M. arvalis* (MicArv1.0) to that of two rodent species for which reference genomes were available: the prairie vole *Microtus ochrogaster* (MicOch1.0) and the house mouse *Mus musculus* (GRCm38.p6). We present in [supplementary table S1, Supplementary Material](#) online assembly metrics for these three species. Overall, the quality of the novel *M. arvalis* genome assembly is very high along the genome: the proportion of ambiguous nucleotides (Ns) is similar among scaffolds (with low values ranging from 2.4% to 5.5%, [supplementary table S2, Supplementary Material](#) online), and the average mappability per scaffold ranged from 84% to 90% ([supplementary table S2, Supplementary Material](#) online). Even though a few regions

in the genome had lower mapping quality, which are likely highly repetitive regions in the chromosomes (e.g. possibly the centromere region on scaffold 1, or telomere at the end of scaffold 22, Fig. 1), different quality metrics such as coverage and mappability remained good all along the genome (Fig. 1). BUSCO (Simão et al. 2015) searches for 303 conserved eukaryotic orthologs as an assessment of the quality of our assembly identified 83.8% of the orthologs in a complete state (compared to 91.1% for the *M. ochrogaster* and 90.8% for the *Mus musculus* genome) and 3.0% in a fragmented state (*M. ochrogaster*: 2.3%; *Mus musculus*: 3.3%). About 13.2% of the BUSCO orthologs are missing (*M. ochrogaster*: 6.6%; *Mus musculus*: 5.9%).

Analysis of syntenic blocks detected numerous chromosomal rearrangements between rodent species (supplementary fig. S2, Supplementary Material online). There were chromosome fusion/fission events (e.g. *M. arvalis* 4th scaffold corresponds to chromosomes 6 and 8 of *M. ochrogaster*, supplementary fig. S2, Supplementary Material online) but also evidence of large inversions (e.g. between scaffold 9 of *M. arvalis* and chromosome 17 of *M. ochrogaster*, supplementary fig. S2, Supplementary Material online). Finally, some regions seemed to contain more rearrangements between the three species than others, such as the center of scaffold 5 in *M. arvalis* (supplementary fig. S2, Supplementary Material online).

Genomic Evolution of the *Microtus* Genus

To study the genome-wide patterns of diversity and divergence between different *Microtus* vole species, we sequenced 12 individuals from 10 different vole species (see Methods, supplementary table S3, Supplementary Material online) that we aligned to the reference *M. arvalis* individual. In addition to the sample used to assemble the reference, our data set included three other *M. arvalis* individuals as well as samples of eight species from the *Microtus* genus and one *Myodes* (*Clethrionomys*) *glareolus* from a related genus (supplementary table S3, Supplementary Material online). The four *M. arvalis* individuals belonged to four previously described evolutionary lineages in Europe potentially undergoing speciation processes: the Central lineage (C), the Western lineage (W), the Eastern lineage (E), and the Italian lineage (I) (Beysard and Heckel 2014; Lischer et al. 2014; Saxenhofer et al. 2019). We obtained a mean sequence depth of about 20× (the depth mode varied from 8 to 28× depending on the sample, supplementary fig. S3 and table S4, Supplementary Material online).

Broad Patterns of Diversity and Divergence

Reference-based alignment and SNP calling identified 75,413,767 polymorphic sites among all 13 newly sequenced vole genomes (see Materials and Methods). The overall levels of genomic diversity assessed from nucleotide diversity (Nei and Li 1979) differed widely between species (Fig. 2a and supplementary fig. S4, Supplementary Material online). The *My. glareolus* genome showed about twice as much genetic diversity as any *Microtus* species (Fig. 2a and supplementary fig. S4, Supplementary Material online). Contrastingly, *M. levis*, which is *M. arvalis*' sister species, showed the lowest overall diversity.

We then used the Pairwise Sequentially Markovian Coalescent (PSMC) (Li and Durbin 2011) to infer changes in effective population size over time. The inference suggested that many of the vole species with suitable data showed a recent decrease in their population size (starting

around 30 to 70 kya depending on the species, Fig. 2b and supplementary fig. S5, Supplementary Material online). Populations of *M. arvalis* were potentially largest before the Last Glacial Maximum, around 30 kya, and declined since then except for the Eastern lineage (Fig. 2b). However, the recent decrease in population size inferred by PSMC corresponds to a recent increase in rate of coalescence, which may be linked to the population structure of *Microtus* species (Mazet et al. 2016). Indeed, if we assume that individuals are taken from subdivided metapopulations, a progressive decrease in coalescence rate going backward in time could simply be due to the scattering (Wakeley 2001) of lineages in different surrounding demes (Mazet et al. 2016; Chikhi et al. 2018). While only intergenic regions (see Materials and Methods) were used in this analysis to mitigate the potential effects of nondemographic processes such as background selection (Johri et al. 2022), masking genic and flanking regions may affect the accuracy of the estimates (Nadachowska-Brzyska et al. 2016; Mather et al. 2020). For example, the high variance in population size estimates observed in recent times (after 5 kya, supplementary fig. S5, Supplementary Material online) are likely artifacts resulting from reduced power in the estimation of N_e (see Wang et al. 2023 for comparison).

We inferred the species' phylogenetic relationships (Fig. 2c) via maximum-likelihood using RAxML (Stamatakis 2014). As expected, genomic divergence within the *Microtus* genus was relatively shallow compared to the outgroup *My. glareolus* but consistent with the expected taxonomic relationships (Jaarola et al. 2004; Fink et al. 2010; Martínková and Moravec 2012; Barbosa et al. 2018), e.g. evolutionary lineages within *M. arvalis* clustering together or sibling species *M. lusitanicus*–*M. duodecimcostatus* in sister relationship).

Identification of Fast Evolving Genomic Regions in Each Species

In order to detect genomic regions potentially involved in adaptations linked to diversification, we identified genomic regions that have been recently evolving rapidly in each species. To do this, we used a novel statistic (RND_{sp}) similar to the PBS statistic (Yi et al. 2010) classically used in genome scans, except that it is based on molecular divergence (d_{xy}) instead of F_{ST} . RND_{sp} measures the relative length of the terminal branch leading to each species since its common ancestor with its closest sister species (based on the topology shown in Fig. 2c). Note that we scaled RND_{sp} relative to the average divergence to the outgroup, *My. glareolus*, in order to correct d_{xy} for its sensitivity to variations in mutation rate along the genome (Burri 2017) (see Materials and Methods). Note also that RND_{sp} absolute values cannot be compared between species, but they are meaningful to evidence fast-evolving regions within each species.

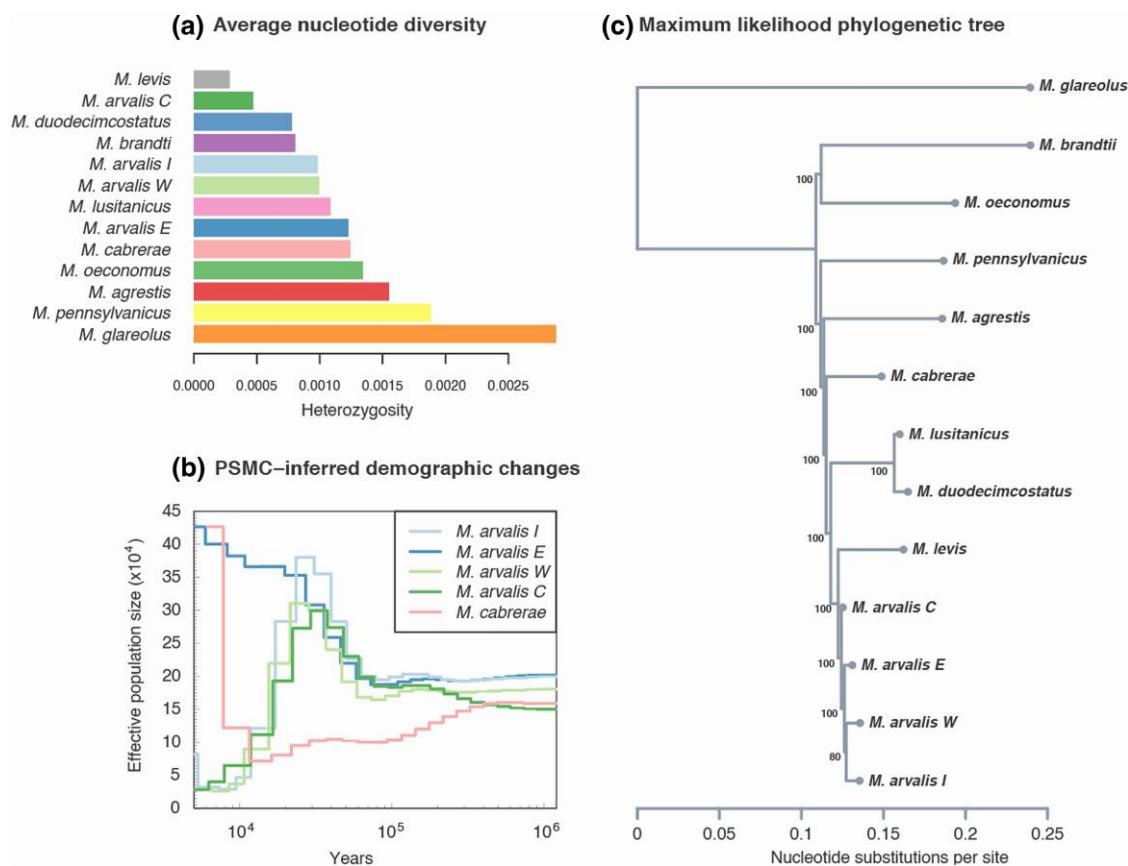


Fig. 2. Broad patterns of diversity and divergence of resequenced individuals. Panel a) corresponds to the average nucleotide diversity (shown as heterozygosity) value per individual. PSMC-inferred effective population size evolution of five samples including the 4 *M. arvalis* lineages is also represented b). This population size estimated assumes a mutation rate of 8.7×10^{-9} per base pair per generation and a generation time of 0.5 yr. Finally, in panel c) we show the maximum-likelihood phylogenetic tree obtained with RAxML. The number of bootstrap replicates supporting this tree topology out of 100 is indicated next to each ancestral node.

Interestingly, we observed systematic variations of RND_{sp} along each scaffold in species with short terminal branches (such as in *M. arvalis* individuals, Fig. 3). This is particularly striking around centromeric regions where we observed slower evolution, potentially due to the impact of background selection in these low recombination regions. When considering a larger evolutionary time scale (i.e. in species with longer terminal branch length, supplementary figs. S6 and S7, Supplementary Material online), the genomic distribution of RND_{sp} is smoother, which is consistent with a main effect of background selection over short time periods (Charlesworth 2012; Nicolaisen and Desai 2013) and the dilution of its effect over longer evolutionary times. Alternatively, the smoothing could be due to a change in the recombination landscapes of species very divergent from the reference genome.

Functional Associations of the Vole Radiation

For each species, we identified the genes found in regions associated with the highest 1% RND_{sp} values, which should

be enriched for fast diverging and potentially adaptive genes. Among the four *M. arvalis* lineages undergoing incipient speciation, we identified many olfactory receptors (OR) among outlier genes (Fig. 3). In *M. arvalis C* lineage, we found eight such receptors on three different scaffolds (Olfr56, and two clusters: Olfr147-149-158-1509 and Olfr497-464-510). In *M. arvalis W* lineage, we identified six other ORs on three separate scaffolds (Olfr18-19-867, Olfr140, Olfr688, and Or51a2). In *M. arvalis I* lineage, we found a single cluster of four highly divergent ORs (Olfr1019-1030-1044-1052). Finally, in *M. arvalis E* lineage, three genomic islands of divergence contain ORs on three different scaffolds (Olfr15-1019-1020-1030, Olfr183-186, and Olfr1509). We note that outlier ORs were not the same in different *M. arvalis* lineages, suggesting that adaptive processes in closely related but parapatric lineages have taken different routes. Other genes commonly represented in regions of high divergence among *M. arvalis* lineages were involved in immunity. This is the case of H2-e genes (Histocompatibility-2, also known as Major Histocompatibility Complex [MHC] class II genes), which

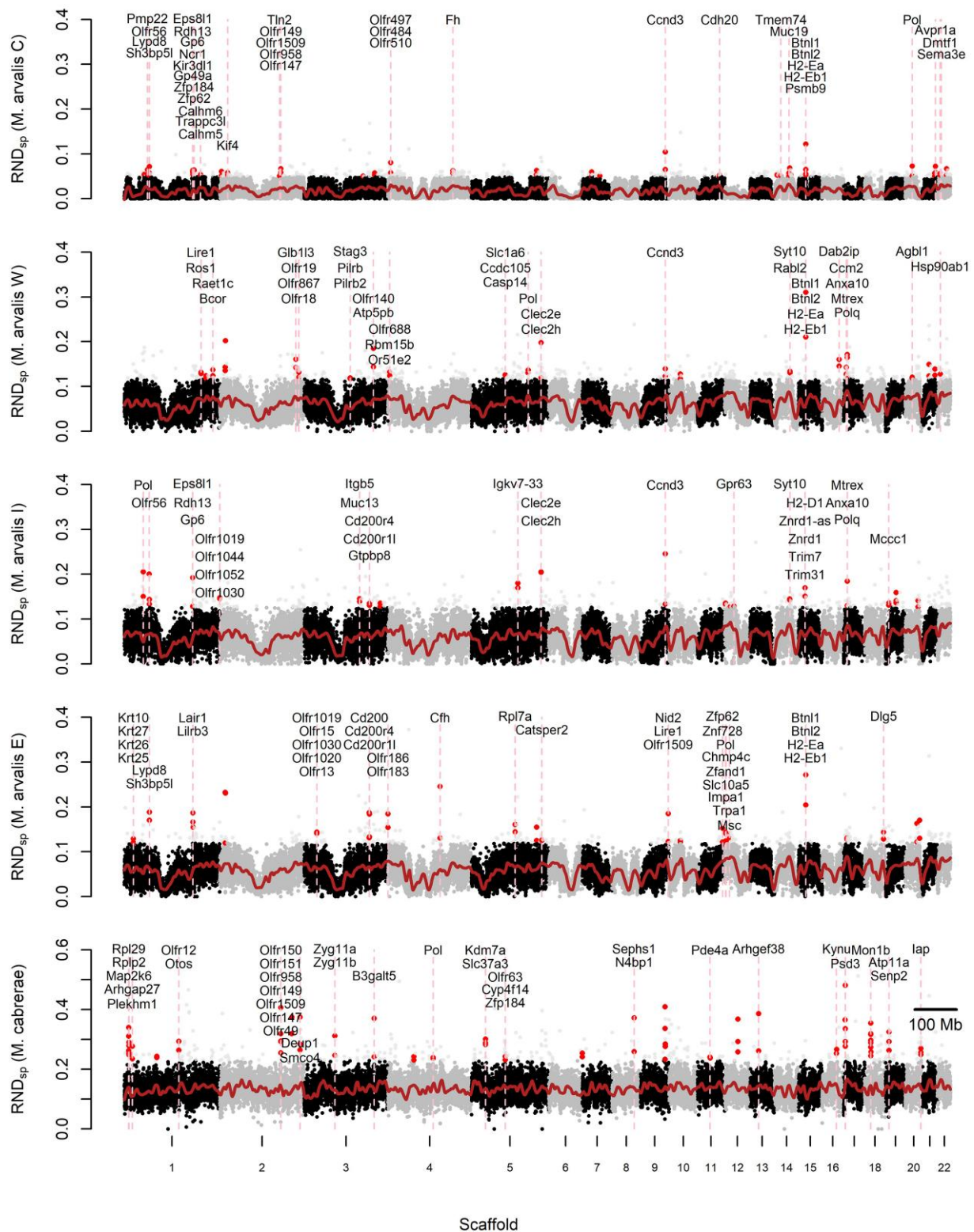


Fig. 3. Genome scan for regions of high divergence in *M. arvalis* individuals and *M. cabreræ* (for other species, see supplementary fig. S6, Supplementary Material online and S7). For each sample, RND_{sp} values computed along 50 kb nonoverlapping windows are represented along the 22 largest scaffolds (gray and black colors indicate the alternance between scaffolds) against its closest relative which has the lowest d_{xy} . The red line corresponds to a loess smoothing of RND_{sp} values. Red dots identify 50 kb windows found in stretches where at least two out of three consecutive windows have an RND_{sp} value in the top 1% of all values. Genes overlapping with such windows are mentioned along vertical pink lines.

were an outlier in two *M. arvalis* lineages (Fig. 3). Other fast-evolving genes involved in immune response included CD200, CD200r4, CD200r1 l, Lypd8, Cfh, and Igkv7-33. Finally, other genes that have been linked to specific biological functions included Ros1 and Catsper2 (both involved in fertility), Avpr1a involved in memory and olfactory learning, as well as a cluster of genes coding for keratins (Krt10-25-26-17).

Genes found in regions of high divergence in *M. levis*, the sister species of *M. arvalis*, also included olfactory receptors (eight genes in three clusters on three different scaffolds, [supplementary fig. S6, Supplementary Material](#) online). We also found a spermatogenesis-associated protein (Spata31), a cluster of three genes involved in development (Foxo1, Slco1a4, and lapp), and two protocadherin-beta genes (Pcdhb14 and Pcdhb18) that are involved in the establishment of interneural connections ([supplementary fig. S6, Supplementary Material](#) online). Highly divergent regions for the two sister species from the Iberian Peninsula, *M. lusitanicus* and *M. duodecimcostatus*, contained genes involved in functions similar to those identified in *M. arvalis* and *M. levis*, i.e. several clusters of olfactory receptors, MHC genes such as H2-Ea and H2-Eb1, other immunity-related genes like Clec2e and Clec2h (C-lectin type receptors) as well as Skint1-2-7, a cluster of genes specifically involved in skin immunity ([supplementary fig. S6, Supplementary Material](#) online). Other outlier genes have a specific role in the neuronal system: the Doublecortin domain-containing protein 2 (Dcdc2), and a cluster of three protocadherin-alpha (Pchda7-4-10) genes coding for neural adhesion proteins, playing a role in the establishment of specific cellular connections in the brain ([supplementary fig. S6, Supplementary Material](#) online). In *M. cabrae*, we also identified highly divergent ORs in three clusters as well as highly divergent metabolism-related genes such as B3galt5, Kynu, or Atp11a (Fig. 3).

We also identified several OR clusters among more divergent species: four clusters in *M. agrestis* (Olf1019-1044-1052-1030, Olf24-15, Olf150-1537-148-958-149, and Olf56), one cluster in *M. brandtii* (Olf1086-1020-481-491-497-510-484), and four clusters in *M. pennsylvanicus* (Olf148-958-149, Olf147, Olf18, and Olf1509) ([supplementary fig. S7, Supplementary Material](#) online). Immunity-related genes are also fast evolving in these species: they include Cdk5rap2 in *M. agrestis*, C-type lectin receptors (Clec2e and Clec2h in both *M. agrestis* and *M. pennsylvanicus*) that are involved in innate immunity ([supplementary fig. S7, Supplementary Material](#) online), Cd200r4 and Cd200r1l in *M. pennsylvanicus*. In *M. agrestis*, genes involved in immunity included Cfh (Complement factor H) and Gbp4 (Guanylate Binding Protein 4). Other outlier genes in this species are involved in metabolism, such as Cyp3a13 (member of the cytochrome P450 family), the serine protease 2 (Prss2), and

the Kynurenine Aminotransferase 3 (Kyt3). The further species had also outlier genes involved in metabolic functions, such as the extended synaptotagmin-1 (Esyn1) in *M. oeconomus*; an aldehyde dehydrogenase (Aldh112) in *M. brandtii*; and a cluster of three Glutathione S-transferase A (Gsta1-2-4) in *M. pennsylvanicus* ([supplementary fig. S7, Supplementary Material](#) online). Finally, other outlier genes with specific functions are involved in cell proliferation and development (ErbB3, Foxo1) in *M. oeconomus*, and are sometimes tissue-specific, such as melanoregulin (Mreg) in *M. pennsylvanicus*, specifically expressed in melanocytes ([supplementary fig. S7, Supplementary Material](#) online).

Gene Ontology Analysis

Gene Ontology (GO) enrichment tests allowed us to identify biological functions that were overrepresented among the top 1% divergent regions in each vole genome ([supplementary table S5, Supplementary Material](#) online). Significant GO terms (FDR < 0.05, [supplementary table S5, Supplementary Material](#) online) found in at least two species are shown in Fig. 4. In line with our results above, many GO terms related to olfaction (e.g. *sensory perception of smell*, *detection of chemical stimulus involved in sensory perception*, *G-protein coupled receptor signaling pathway*) were significant after correction for multiple testing in every species we studied. Two other groups of GO terms were recurrently enriched in several species, although at a lesser extent than olfaction-related genes. These are immunity-related GO terms, such as *immune response* in most species, *innate immune response* in *M. lusitanicus*, *M. agrestis*, and *M. pennsylvanicus* or MHC GO terms (e.g. *antigen processing and presentation of peptide antigen via MHC class I* in *M. levis*, *M. lusitanicus*, and *M. duodecimcostatus*). The second group contained GO terms linked to metabolism (particularly lipid metabolism, e.g. *fatty acid derivative metabolic process* in *M. oeconomus*, *M. cabrae*, and *M. pennsylvanicus* or *oxido-reduction process* in *M. cabrae*, *M. agrestis*, and *M. oeconomus*). The strong representation of olfaction-related terms was confirmed even when performing the test on a conservative restricted set of genes found in islands of divergence, i.e. in regions where at least 150/250 kb were significantly divergent (see Materials and Methods). In that case, olfaction-related GO terms remained significant in 10 out of 12 species ([supplementary fig. S8, Supplementary Material](#) online).

Highly Divergent Olfaction-related Genes Show Little Diversity Within Species

As ORs in vole species showed exceptional patterns of divergence and fast evolution, we more specifically investigated the species distribution and genetic diversity of olfaction-related genes in the top 1% divergent regions



Fig. 4. GO terms enriched for highly divergent genes found in more than one *Microtus* vole species. Each dot represents a significant GO term (FDR < 0.05) in a given species. Dot size is proportional to the number of outlier genes within this GO term, and darker colors indicate more significant enrichment tests (lower q-value). Enrichment tests have been performed using all genes found in the top 1% divergent genomic windows (see Materials and Methods). Each row is labeled with the GO term description and numbers between parentheses indicating i) the number of highly divergent genes that are specific to any genome, ii) is total number of highly divergent genes in this GO term, and iii) the total number of genes in this GO term that have been tested.

(Fig. 5a). As olfaction-related genes, we considered here both ORs (expressed in the olfactory epithelium) and vomeronasal receptors (expressed in the vomeronasal organ). Many of these genes were outliers in several vole species. Interestingly, most of these outlier genes presented very low diversity in the species where they were outlier (Figs. 5a and 6a, 6b), suggesting that divergence is driven by the fixation of beneficial mutations during selective sweeps. Note however that a very few divergent genes presented both high and low diversity patterns within focal species, such as Olfr49, Olfr1019, or Olfr1020 (Fig. 5a). In order to assess more quantitatively how the intraspecific diversity of ORs compares to other genes, we compared the distribution quantiles of diversity values for different sets of genes:

outlier ORs, outlier genes that are not ORs, and randomly sampled genes (supplementary figs. S9 and S10, Supplementary Material online). We found that, in all species, levels of intraspecific diversity were lower in highly divergent regions containing ORs than in randomly selected regions. OR outlier regions had also lower levels of diversity than non-OR outliers in most species, except in *M. oeconomus* and *M. levis*, where ORs and non-OR outliers had similarly low levels of diversity, and in *M. brandtii* and *M. agrestis* where ORs had higher diversity than other outlier genes (Fig. 5, supplementary figs. S9 and S10, Supplementary Material online). Most outlier genes including ORs did not show exceptionally high heterozygosity but rather fell into the distribution of the genomic background

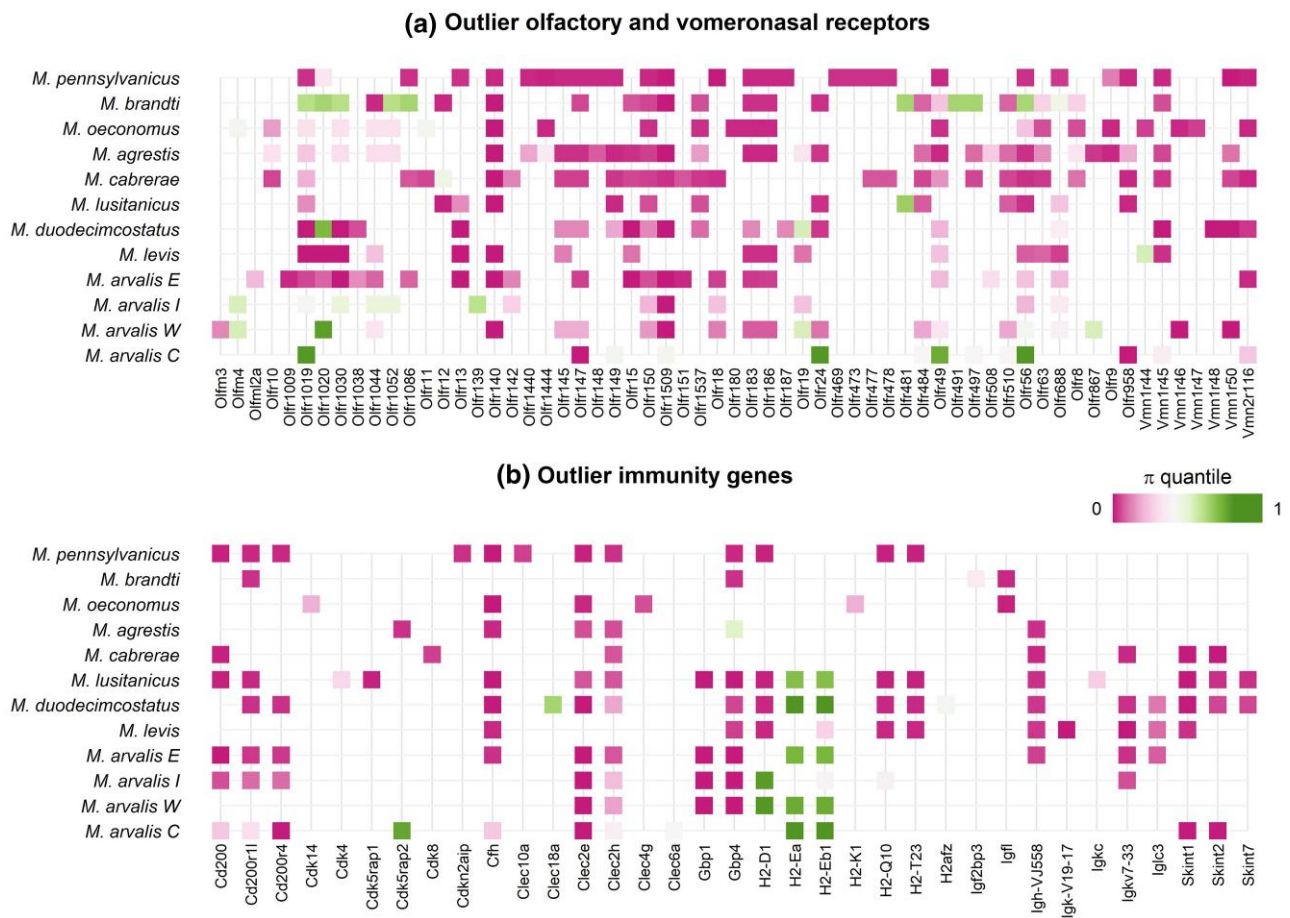


Fig. 5. Patterns of diversity at a) olfactory and vomeronasal receptors and b) immunity genes found in the top 1% divergent windows of multiple or single *Microtus* species. The color of the square corresponds to the nucleotide diversity of the outlier window, expressed as the quantile of the focal genome’s diversity distribution (low diversity is pink, high diversity is green).

(supplementary fig. S11, Supplementary Material online), which supports that these diversity patterns were little impacted by the potential mapping of paralogous OR sequences.

We performed a similar analysis on immune-related genes (Fig. 5b) present in our top 1% highly divergent regions. While the most divergent genes presented reduced diversity relative to other genes in the genome (see e.g. gene Cd200 in Fig. 6d), two histocompatibility genes (H2-Ea and H2-Eb1) clearly showed higher diversity than average (as shown in Fig. 6c), compatible with some form of balancing selection.

Discussion

The chromosome-scale assembly of a vole reference genome and the sequencing of additional genomes representing the explosive radiation in the *Microtus* genus has enabled us to identify several evolutionary processes advancing divergence among species. Using whole genomes of 12 individuals from 10 different species, we showed that

olfaction played a particular role in the genomic diversification of *Microtus*, but that immunity and metabolic processes are also unusually fast evolving in most studied species. This suggests that adaptation to environmental pressure but also sexual selection and kin recognition may have important roles in the radiation of voles.

Fast Evolution of the Olfactory Pathway in the Vole Radiation

Sensory systems play a fundamental role in species isolation, particularly through behavioral processes (Smadja and Butlin 2009). Visual cues are often positively selected in diverse taxa, leading for example to the particular evolution of opsin genes in fish (Gaither et al. 2015), wing pigmentation in *Heliconius* butterflies (Martin et al. 2013; Mazo-Vargas et al. 2017), or an enormous color diversity in cichlids (Meier et al. 2018). Chemosensory systems are also commonly involved in behavioral isolation of many taxa, particularly in insects such as moths and fruit flies but also in rodents (Smadja and Butlin 2009). For example,

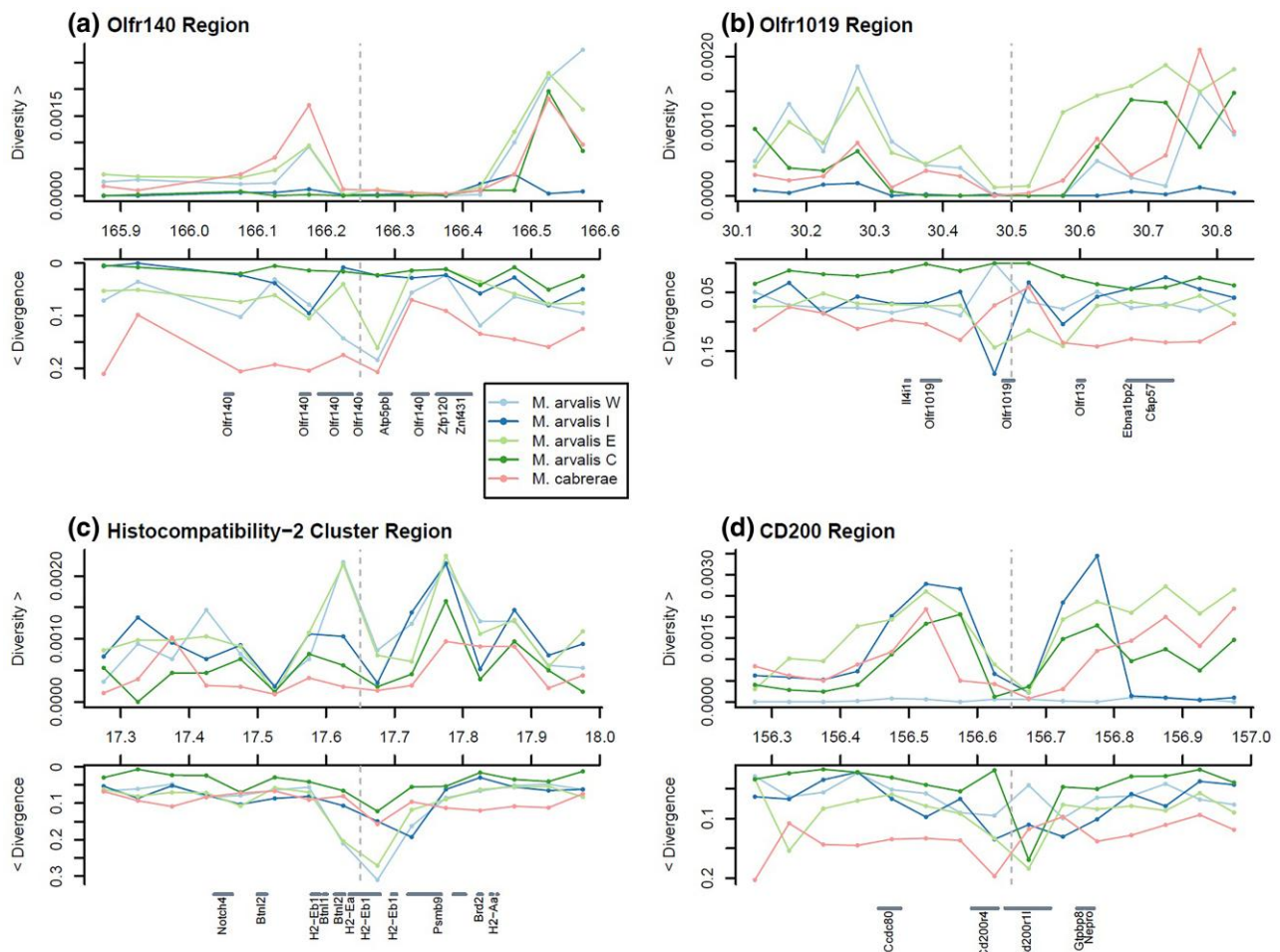


Fig. 6. Zooms on patterns of diversity and divergence among three of the most significantly divergent windows in *M. arvalis* individuals and *M. cabreræ*: a) the Olfr140 region on scaffold 3, b) the Olfr1019 region on scaffold 3 c) the Histocompatibility-2 (H2) cluster region on scaffold 15 and d) the CD200 region on scaffold 3. For each of the five selected individuals (four *M. arvalis* lineages and *M. cabreræ*), window-based average nucleotide diversity (π) and RND_{sp} are represented along a 1 Mb region surrounding the most divergent window on average (indicated with a dashed gray vertical line).

studies of hybrid zones revealed the importance of olfactory cues in mate preference in the house mouse (Smadja et al. 2004, 2015) or pine voles (Cerveira et al. 2019). In *Microtus*, partial or complete reproductive isolation has been shown to depend on behavioral factors, and olfaction can play a key role in mate recognition (Soares 2013; Duarte et al. 2016; Cerveira et al. 2019). Given the absence of major morphological differentiation and often subterranean lifestyles in voles, olfaction may play an essential role in the detection and maintenance of species differences in *Microtus* species, e.g. through mate choice (Smadja and Butlin 2009).

Our study shows that the fast *Microtus* radiation was accompanied by particularly rapid evolution of genes in the olfactory pathway. Overall, fast-evolving genes were involved in different stages of the olfactory response: odorant production (e.g. pheromones), reception (through

olfactory receptor neurons), and the downstream processing of olfactory information (e.g. through the olfactory bulb, memory and behavioral response). Olfaction-related genes are among the most divergent regions of the genome in all species we studied (Fig. 3; supplementary figs. S6 and S7, Supplementary Material online), and some of these genes such as ORs and vomeronasal receptors are outliers in multiple species (Fig. 6). It is important to note that the deeper evolutionary history of ORs in mammals is characterized by extensive gene duplication and reduction (Hughes et al. 2018), and it is plausible that this has continued within the *Microtus* radiation. Even though some of the ORs may not be orthologous in all the analyzed *Microtus* species, generally low diversity and nonelevated heterozygosity of outlier genes (Figs. 5 and supplementary fig. S9 to S11, Supplementary Material online) suggests that their detection as fast evolving was not driven by erroneous

mapping of paralogous sequences. Of course, high rates of potential gene expansion or loss in the history of diversification would also be consistent with fast evolution of ORs in the *Microtus* genomes. Fast evolution of genes can be a result of either adaptation or relaxed purifying selection under suitable conditions, which has been found especially in island populations of birds and rodents (Bliard et al. 2020; Wang and Heckel 2024). However, for our analyses spanning multiple species, relaxation of selection would have needed to occur in continental populations of different species with overall very high effective population sizes, so it is more parsimonious that the observed pattern of fast evolution is a result of adaptation.

ORs are not the only olfaction-related genes presenting high levels of divergence. We also identified several other genes that could be directly involved in the production and the decoding of olfactory signals. For instance, some nonvolatile pheromones known to be involved in social recognition in mice are peptide ligands of the MHC encoded by the Histocompatibility-2 (H2) genes, which are outliers in multiple species. Other types of pheromones can be the target of selection in some species: *Kyat3*, which is an outlier gene in *M. agrestis*, encodes for an enzyme involved in the metabolism of the kynurenine pheromone. Finally, some genes that are involved downstream in the olfactory pathway through neuronal processes are among the most divergent genes. As an example, some proteins are involved in establishing specific cellular connections in the brain and play a role in regulating learning and memory functions (e.g. *Dcdc2*, the protocadherin-alpha cluster, or *Avpr1a*). Knock-out experiments on mice have revealed the impact on learning and behavioral response for these genes (Caldwell et al. 2008).

Signatures of Environmental Pressures in Patterns of Divergence

Ecological adaptations are also central to the formation of new species when reproductive isolation evolves as a result of divergent selection (Rundle and Nosil 2005). Besides sexual selection, other selective pressures could induce a response in olfaction-related traits, for example habitat and diet preferences (Hughes et al. 2018; Auer et al. 2020), or predator sensing. The only gene found to be an outlier in all the species, *Olfr1019*, is known for its important role in the behavioral reactions to a fox urine odorant, 2,4,5-trimethylthiazoline (Saito et al. 2017), which elicits strong fear responses in mice (Endres and Fendt 2009). This suggests a strong selective pressure of innate odor response on voles which are preyed upon by a wide range of mammalian predators (Klemola et al. 1997). We detected also two divergent genes that are involved in positive response to odorants commonly found in plants (de March, et al. 2020): *Olfr19* in the response to whiskey lactone,

common in oak trees (Doussot et al. 2002); *Olfr183* to isoamyl acetate, naturally produced by ripening fruit. As voles are largely herbivorous, they might present specific adaptations to diet as they depend on the availability of annual plants, often grasses, throughout the year. Even though voles prefer feeding on annual plants, they also consume perennials when the ecological conditions require it (Rosário et al. 2008). As slow-growing plants typically contain more toxic secondary metabolites, being able to detoxify these molecules could confer a strong selective advantage (Francis et al. 2005; Bock 2016; Johnson et al. 2018). This could explain the presence of metabolism-related genes, particularly detoxification enzymes such as Glutathione S-transferase A or cytochrome P450, among highly divergent windows. Metabolism-related genes might also play a role in adaptation to various temperature conditions as *Microtus* species cover vast latitude and altitude ranges (Carleton and Musser 2005). Thermal adaptation can occur through reduced metabolic rate (Bozinovic et al. 2005), or changes in energy metabolism and stress response (Quina et al. 2015; Monarca et al. 2019).

In addition, several genes and GO terms related to immunity were observed among divergent genomic windows in most species. The evolution of the immune system is often the consequence of host–pathogens interactions and arms races, leading to diversifying selection and the maintenance of polymorphisms. Long-term balancing selection has been inferred to act on immunity-related genes in various organisms, for example in primates (Maibach et al. 2017; Bitarello et al. 2018), *Drosophila* (Croze et al. 2017; Chapman et al. 2019), or birds (O'Connor et al. 2016). In voles, immunity-related genes (e.g. H2, *Skint*, and *Clec* families) are commonly found in highly divergent genomic regions. H2 genes (Histocompatibility-2) are part of the MHC that is common to all jawed vertebrates (Klein 1986). Beyond their role in response to pathogens, MHC genes are also potentially involved in sexual selection and kin recognition through mating preferences that promote the MHC diversity of offspring (Eizaguirre et al. 2009). We identified further immunity-related gene families that have been recently fast evolving in multiple species, such as the *Skint* gene family, which is likely unique to rodents (Sutoh et al. 2018), C-Type Lectin-Like Receptors, Immunoglobulins and 2'-5'-oligoadenylate synthetase genes.

Conclusion

Overall, our results show the fast evolution of genes related to the olfactory system, as well as immunity during vole species' divergence, indicating the importance of selection in related traits for the radiation of the group. The striking morphological similarity in this highly speciose genus suggests that conspecific identification occurs through other

channels than visual cues, and our analyses indicate that olfactory signals are among the key elements for this discrimination. These observations open several questions and perspectives for this study system. Future directions could, for example, include a more detailed characterization of the full repertoire of ORs in *Microtus* and their roles in the recognition of con- or heterospecific individuals or kin, functional validation of the importance of some candidate genes through experimental studies, or investigating the diversity of odorant molecules. Additional genomic data of several individuals per species could provide further insights in differences in OR diversity within and between species to potentially characterize their role in ecological and evolutionary interactions. At last, our reference genome should facilitate the collection and genomic analysis of large samples that would allow one to better investigate patterns of gene flow between closely related rodent taxa (e.g. Saxenhofer et al. 2022; Labutin and Heckel 2024), and to better understand the complex evolutionary processes associated with vole species divergence (e.g. Baca et al. 2023; Wang et al. 2023).

Materials and Methods

Microtus arvalis Genome Assembly

Assembly Generation

DNA was extracted from a male *M. arvalis* individual, the first-generation offspring of a cross between individuals sampled from the wild (Bize et al. 2018). DNA extraction and sequencing library preparation was performed by Dovetail Genomics (Santa Cruz, CA). Extracted DNA was used to produce long-range sequencing libraries using the “Chicago” (Putnam et al. 2016) and Hi-C (Belton et al. 2012) methods by Dovetail Genomics (Santa Cruz, CA). The *M. arvalis* genome was assembled in three phases: (i) a *de novo* assembly was generated based on 150 bp paired-end Illumina reads; (ii) the *de novo* assembly was scaffolded by Chicago libraries; (iii) the Chicago assembly was scaffolded by Hi-C libraries. Hi-C libraries allow the scaffolding to operate at a larger scale, potentially at the chromosome level. Additional RNAseq for genome annotation was performed by the Next Generation Sequencing Platform of the University of Bern using equimolar libraries of several tissues (muscle, liver, spleen, brain, and testes) from the same individual resulting in 302,714,054 paired-end Illumina reads (2×150 bp).

Assembly Quality Assessment

Basic assembly metrics were computed using a script from assemblathon.org (<http://assemblathon.org/post/44431933387/assemblathon-2-basic-assembly-metrics>). The metrics NG50 and LG50 assume a genome size of 2.54 gigabases, which was estimated using flow cytometry (Hare

and Johnston 2011). The quality and completeness of the assembly were assessed by searching for universal single-copy orthologues implemented in *BUSCO* version 2.0 (Simão et al. 2015) using the Eukaryota database containing 303 orthologues. These numbers were put in relation to the same *BUSCO* analysis run on the assemblies of *M. ochrogaster* and *Mus musculus*. The quality and completeness of the assembly were also inferred by mapping the Illumina raw reads used to assemble the genome back to the assembly. The mapper *bowtie2* version 2.2.4 (Langmead and Salzberg 2012) was used with default parameters and a maximum allowed insertion size of 1000 bp, to exclude the presence of large mis-assemblies. We further assessed the quality and completeness of the assembly using RNAseq data. The RNAseq libraries were mapped to the genome using *STAR* version 2.5.3a (Dobin et al. 2013) with default parameters. The evaluation of repeats and low complexity regions were assessed using *RepeatModeler* version 1.0.8 (Smit and Hubley 2008) and *RepeatMasker* version 4.0.7 (Smit et al. 2015) with default parameters.

Gene Prediction and Annotation

A model for gene prediction was trained as described in Tran et al. (2016). RNAseq expression data were mapped onto the assembly using *STAR* version 2.5.3a (Dobin et al. 2013) and transcripts obtained with *cufflinks* version 2.2.1 (Trapnell et al. 2012). Transcripts were translated into amino acid sequences using *transeq* from *EMBOSS* version 6.5.7 (Rice et al. 2000). Uniprot *Mus musculus* reference proteins (version: September 2017) were mapped with *gsearch36* version 36.3.5e (Pearson 1991) onto the translated sequences to identify the right coding phase, start and stop-position. High confidence CDS were translated into an *Augustus* compatible format and used as a training set to build the gene-model. Genes were predicted using *Augustus* version 3.0.2 (Stanke and Waack 2003) with the following parameters (`-strand=both` `-genemodel=partial` `-alternatives-from-sampling=true`). Predicted genes were annotated based on homology to the uniprot *Mus musculus* reference proteome using the *BLASTP* with default parameters. For each predicted gene, the predicted homolog with the lowest e-value was kept.

Synteny Analysis

We computed the pairwise synteny between the *M. arvalis* genome, *M. ochrogaster* (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000317375.1/) and *Mus musculus* (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001635.26/) assemblies using the online tool *D-Genies* (Cabanettes and Klopp 2018) with default parameters. Plots

were created with custom scripts in R using packages *GenomicRanges*, *bezier*, and *ggsci*.

Vole Resequencing Data Generation and Analysis

Samples and Sequencing Library Preparation

We resequenced 12 individuals from 10 different vole species from Europe, Asia and America that are described in [supplementary table S2, Supplementary Material](#) online. Separate DNA libraries were prepared for each individual by the Next-Generation Sequencing Platform of the University of Bern using the Illumina TruSeq DNA PCR-Free kit following the recommendation of the manufacturer. Libraries were sequenced at the Next-Generation Sequencing Platform of the University of Bern on an Illumina HiSeq 3000 machine (2 × 150 bp). We obtained between 82 and 226 (mean: 159) million sequence reads per sample with mean quality scores between 36.4 and 37.8.

Variant Calling

Paired-end reads were first aligned to the *M. arvalis* reference genome with *stampy* version 1.0.32 (Lunter and Goodson 2011), using a substitution rate of 0.05. We then realigned reads around indels using GATK version 3.8-0-ge9d806836 (McKenna et al. 2010). The proportion of aligned reads ranged from 0.85 in *My. glareolus* to 0.95 in the three *M. arvalis* samples ([supplementary table S3, Supplementary Material](#) online). Variant calling was performed using *GATK UnifiedGenotyper* without PCR correction and with a 0.001 heterozygous prior (allowing for equal priors for both homozygous calls, `-pcr_error_rate 0 -inputPrior 0.001 -inputPrior 0.4995`). The mode of the depth of sequence coverage distributions ranged from 8 to 28, with an average of 17.25 excluding the reference MarvC ([supplementary table S3, Supplementary Material](#) online). Given the differences in depth distributions ([supplementary fig. S2, Supplementary Material](#) online), we defined custom depth filters for each individual. We then filtered called genotypes based on the depth of coverage in the tails of the distributions using custom DP filters for each sample (2 SD around the mode were kept, [supplementary fig. S2 and table S3, Supplementary Material](#) online). We kept only diallelic sites and filtered out genotypes based on their quality ($GQ \leq 30$ for all samples) and the mapping quality ($MQ < 25$; $MQ0 \geq 4$; $[MQ0/DP] > 0.1$). We kept sites with no missing genotypes. After filtering, our dataset consisted in 75,413,767 polymorphic sites and 679,864,446 invariant sites.

Phylogeny

To infer the evolutionary relationships of the 13 sequenced individuals, we generated a PHYLIP alignment from the

previously generated VCF file using the *vcf2phylip* v1.5 Python script (<https://github.com/edgardomortiz/vcf2phylip>). For heterozygous genotypes, a consensus was made and genotypes were replaced by IUPAC nucleotide ambiguity codes. Species' phylogenetic relationships were inferred with *RAXML* version 8.2.4 (Stamatakis 2014), based on the GTR + GAMMA substitution model and default parameters. The Broyden–Fletcher–Goldfarb–Shanno algorithm was used for GTR model parameters optimization. 100 bootstrap replicates were performed with *RAXML* rapid bootstrapping option. The most likely tree and bootstrap values were then represented using an R custom script (https://github.com/CMPG/vole_genome_analysis).

Demographic Inference

We inferred changes in effective population size using the PSMC model (Li and Durbin 2011) implemented in the *psmc* software (<https://github.com/lh3/psmc>). We generated the *psmc* input file using the provided *fq2psmcfa* script after filtering FASTQ files for read depth between 3 and 120 and quality ≥ 30 . Only intergenic regions which are 10 kb away from any gene were kept. We then ran *psmc* using the following command for each sample: `psmc -N40 -t15 -r3 -p "6*1+20*2+4+6"`. We used a mutation rate of 8.7×10^{-9} and generation time of 0.5 yr following Wang et al. (2023).

Population Genomics Analyses

Different population genomics statistics were computed along the genome using *PopGenome* R package (Pfeifer et al. 2014). Statistics were computed in nonoverlapping windows of 50 kb along the 22 largest scaffolds. We then computed the nucleotide diversity π and the absolute nucleotide divergence d_{xy} (Nei and Li 1979). To do so, we calculated the absolute nucleotide diversity (within and between, respectively) using the *F_ST.stats* method from *PopGenome* package, and divided it by the number of non-missing sites present in a given genomic window. We also computed a lineage-specific index of divergence. In a similar fashion to *PBS* (Yi et al. 2010), we define d_{sp} as the length of the branch leading to a given lineage. It is computed as follows:

$$d_{sp} = \frac{d_{AB} + d_{AO} - d_{BO}}{2},$$

where d_{ij} is the d_{xy} value between species i and j . A and B correspond to two closely related species (or evolutionary lineages in the case of *M. arvalis*) which have the lowest d_{xy} on average, and O is an outgroup (in our case, it is *My. glareolus* for all *Microtus* species). Then, to have a measure that is relatively insensitive to mutation rate

variation along the genome, we compute a metric similar to the Relative Node Depth *RND* (Feder et al. 2005) but with our lineage-specific estimate of divergence. The d_{sp} statistic rescaled using an outgroup is simply defined as

$$RND_{sp} = \frac{d_{sp}}{\bar{d}_O}$$

where \bar{d}_O is the average absolute divergence of all *Microtus* species to the *My. glareolus* outgroup. This measure is essentially an extension of the *PBS* statistic where the measure of population differentiation F_{ST} is replaced by nucleotide divergence, and is expected to correct for variations in d_{xy} due to mutation rate differences along the genome (Hellmann et al. 2005; Castellano et al. 2020), assuming that these variations in mutation rate are the same in *Myodes* and *Microtus* species. Indeed, we observed very strong correlations in d_{xy} values among comparisons of different species pairs (Spearman's correlation coefficients range from 0.77 to 0.99, [supplementary fig. S12, Supplementary Material](#) online). This observation is in keeping with similar local mutation rates in all species affecting levels of divergence, justifying the normalization in RND_{sp} . Note also that this correlation is found higher for less related species: the greater the average divergence, the greater the correlation between pairwise divergence values ([supplementary fig. S12, Supplementary Material](#) online).

Functional Analysis of Outlier Regions

We mapped outlier genomic windows to predicted genes in the previously generated *M. arvalis* annotations. We kept only genes with at least one homolog in the mouse genome (*Mus musculus*). For each gene, we kept the mouse homolog with the lowest e-value resulting from the BLASTP analysis. Duplication and partial or complete de-functionalization of genes (e.g. for ORs) might lead to different genomic repertoires between the *Microtus* species. Using evolutionarily relatively distant mouse homologs as a common reference may thus miss the fastest evolving regions in some *Microtus* genomes but it is also expected to limit their quantitative impact in the functional association of outlier regions. We mapped genes to genomic windows if they had at least 1 base pair overlap with the 50 kb window of interest.

Each time an outlier list of genes was generated, we performed a Gene Ontology enrichment test using the *STRINGdb* (Franceschini et al. 2012) R package, using GO terms from *Mus musculus*. The background list of genes (i.e. the set of all tested genes) is the list of all genes for which we had a mouse homolog in the annotation (13,543 genes). Enrichment tests were Fisher's exact tests with a False Discovery Rate (FDR) computation following the Benjamini–Hochberg procedure (Benjamini and

Hochberg 1995). We considered GO terms as significant if they had an FDR < 0.05. We filtered out GO terms that were too general by removing GO terms containing more than 1,500 genes in the background list.

Supplementary Material

[Supplementary material](#) is available at *Genome Biology and Evolution* online.

Acknowledgments

We would like to thank A. Thiéry for bioinformatics support, S. Tellenbach for laboratory assistance, and C. Bastos-Silveira, J. Bryja, M. Fischer, Y. Liu, K. Norrdahl, J. Runge, L. Yalkovskaya, and K. Zub for assistance with sample collection. We are grateful for constructive comments and suggestions from the anonymous reviewers. We thank the Next-Generation Sequencing Platform of the University of Bern for high-throughput sequencing. This work was supported by grants 31003A_149585 and 176209 from the Swiss National Science Foundation to G.H. X.W. received a scholarship from the China Scholarship Council (grant no. 201706380049). A.G. was partially supported by a grant no 31003A_143393 to L.E.

Data Availability

The reference genome MicArv1.0 can be found on NCBI RefSeq database (BioProject accession number: PRJNA737461). This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAHRII000000000. The version described in this paper is version JAHRII010000000. All R and Bash scripts used for resequencing data processing, analysis and visualization, as well as preprocessed resequencing data used for population genomics analyses, can be found on GitHub (https://github.com/CMPG/vole_genome_analysis/).

Literature Cited

- Auer TO, Khallaf MA, Silbering AF, Zappia G, Ellis K, Álvarez-Ocaña R, Arguello JR, Hansson BS, Jefferis GSXE, Caron SJC, et al. Olfactory receptor and circuit evolution promote host specialization. *Nature* 2020;579(7799):402–408. <https://doi.org/10.1038/s41586-020-2073-7>.
- Baca M, Popović D, Lemanik A, Bañuls-Cardona S, Conard NJ, Cuenca-Bescós G, Desclaux E, Fewlass H, Garcia JT, Hadravova T, et al. Ancient DNA reveals interstadials as a driver of common vole population dynamics during the last glacial period. *J Biogeogr.* 2023;50(1):183–196. <https://doi.org/10.1111/jbi.14521>.
- Barbosa S, Paupério J, Pavlova SV, Alves PC, Searle JB. The *Microtus* voles: resolving the phylogeny of one of the most speciose mammalian genera using genomics. *Mol Phylogenet Evol.* 2018;125:85–92. <https://doi.org/10.1016/j.ympev.2018.03.017>.
- Bastos-Silveira C, Santos SM, Monarca R, Mathias M, Heckel G. Deep mitochondrial introgression and hybridization among ecologically

- divergent vole species. *Mol Ecol.* 2012;21(21):5309–5323. <https://doi.org/10.1111/mec.12018>.
- Belton J-M, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 2012;58(3):268–276. <https://doi.org/10.1016/j.ymeth.2012.05.001>.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc.* 1995;57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Beysard M, Heckel G. Structure and dynamics of hybrid zones at different stages of speciation in the common vole (*Microtus arvalis*). *Mol Ecol.* 2014;23(3):673–687. <https://doi.org/10.1111/mec.12613>.
- Beysard M, Krebs-Wheaton R, Heckel G. Tracing reinforcement through asymmetrical partner preference in the European common vole *Microtus arvalis*. *BMC Evol Biol.* 2015;15(1):1–8. <https://doi.org/10.1186/s12862-015-0455-5>.
- Beysard M, Perrin N, Jaarola M, Heckel G, Vogel P. Asymmetric and differential gene introgression at a contact zone between two highly divergent lineages of field voles (*Microtus agrestis*). *J Evol Biol.* 2012;25(2):400–408. <https://doi.org/10.1111/j.1420-9101.2011.02432.x>.
- Bitarello BD, De Filippo C, Teixeira JC, Schmidt JM, Kleinert P, Meyer D, Andrés AM. Signatures of long-term balancing selection in human genomes. *Genome Biol Evol.* 2018;10(3):939–955. <https://doi.org/10.1093/gbe/evy054>.
- Bize P, Lowe I, Lehto Hürlimann M, Heckel G. Effects of the mitochondrial and nuclear genomes on nonshivering thermogenesis in a wild derived rodent. *Integr Comp Biol.* 2018;58(3):532–543. <https://doi.org/10.1093/icb/icy072>.
- Bliard L, Paquet M, Robert A, Dufour P, Renoult JP, Grégoire A, Crochet P-A, Covas R, Doutrelant C. Examining the link between relaxed predation and bird coloration on islands. *Biol Lett.* 2020;16(4):20200002. <https://doi.org/10.1098/rsbl.2020.0002>.
- Bock KW. The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: animal-plant arms-race and co-evolution. *Biochem Pharmacol.* 2016;99:11–17. <https://doi.org/10.1016/j.bcp.2015.10.001>.
- Bozinovic F, Carter MJ, Ebensperger LA. A test of the thermal-stress and the cost-of-burrowing hypotheses among populations of the subterranean rodent *Spalacopus cyanus*. *Comp Biochem Physiol A Mol Integr Physiol.* 2005;140(3):329–336. <https://doi.org/10.1016/j.cbpb.2005.01.015>.
- Burri R. Linked selection, demography and the evolution of correlated genomic landscapes in birds and beyond. *Mol Ecol.* 2017;26:3853–3856. <https://doi.org/10.1111/mec.14167>.
- Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* 2018;6:e4958. <https://doi.org/10.7717/peerj.4958>.
- Caldwell HK, Wersinger SR, Young WS III. The role of the vasopressin 1b receptor in aggression and other social behaviours. *Prog Brain Res.* 2008;170:65–72. [https://doi.org/10.1016/S0079-6123\(08\)00406-8](https://doi.org/10.1016/S0079-6123(08)00406-8).
- Carleton MD, Musser GG. Subfamily arvicolinae. In: Wilson DE, Reeder DM, editors. *Mammal species of the world: a taxonomic and geographic reference*. Baltimore: Johns Hopkins University Press; 2005. p. 956–1039.
- Castellano D, Eyre-Walker A, Munch K. Impact of mutation rate and selection at linked sites on DNA variation across the genomes of humans and other homininae. *Genome Biol Evol.* 2020;12(1):3550–3561. <https://doi.org/10.1093/gbe/evz215>.
- Cerveira AM, Soares JA, Bastos-Silveira C, Mathias MDL. Reproductive isolation between sister species of Iberian pine voles, *Microtus duodecimcostatus* and *M. lusitanicus*. *Ethol Ecol Evol.* 2019;31(2):121–139. <https://doi.org/10.1080/03949370.2018.1508075>.
- Chapman JR, Hill T, Unckless RL. Balancing selection drives the maintenance of genetic variation in *Drosophila* antimicrobial peptides. *Genome Biol Evol.* 2019;11(9):2691–2701. <https://doi.org/10.1093/gbe/evz191>.
- Charlesworth B. The effects of deleterious mutations on evolution at linked sites. *Genetics* 2012;190(1):5–22. <https://doi.org/10.1534/genetics.111.134288>.
- Chikhi L, Rodriguez W, Grusea S, Santos P, Boitard S, Mazet O. The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. *Heredity (Edinb).* 2018;120(1):13–24. <https://doi.org/10.1038/s41437-017-0005-6>.
- Comeault AA, Flaxman SM, Riesch R, Curran E, Soria-Carrasco V, Gompert Z, Farkas TE, Muschick M, Parchman TL, Schwander T. Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. *Curr Biol.* 2015;25(15):1975–1981. <https://doi.org/10.1016/j.cub.2015.05.058>.
- Coyne JA, Orr HA. *Speciation*. Sunderland (MA): Sinauer Associates; 2004.
- Croze M, Wollstein A, Božičević V, Živković D, Stephan W, Hutter S. A genome-wide scan for genes under balancing selection in *Drosophila melanogaster*. *BMC Evol Biol.* 2017;17(1):15–12. <https://doi.org/10.1186/s12862-016-0857-z>.
- Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol.* 2014;23(13):3133–3157. <https://doi.org/10.1111/mec.12796>.
- De Jonge G, Ketel NA. An analysis of copulatory behaviour of *Microtus agrestis* and *M. arvalis* in relation to reproductive isolation. *Behaviour* 1981;78(3-4):227–258. <https://doi.org/10.1163/156853981X00338>.
- De March CA, Titlow WB, Sengoku T, Breheny P, Matsunami H, McClintock TS. Modulation of the combinatorial code of odorant receptor response patterns in odorant mixtures. *Mol Cell Neurosci.* 2020;104:103469. <https://doi.org/10.1016/j.mcn.2020.103469>.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
- Dobzhansky TG. *Genetics and the origin of Species*. New York: Columbia University Press; 1937.
- Doussot F, De Jésus B, Quideau S, Pardon P. Extractives content in cooperation oak wood during natural seasoning and toasting; influence of tree species, geographic location, and single-tree effects. *J Agric Food Chem.* 2002;50(21):5955–5961. <https://doi.org/10.1021/jf020494e>.
- Duarte MA, Heckel G, da Luz Mathias M, Bastos-Silveira C. (Duarte2016 co-authors). 2016. Olfactory receptors and behavioural isolation: a study on *Microtus* voles. *Mamm Res.* 2016;61(4):399–407. <https://doi.org/10.1007/s13364-016-0266-0>.
- Eizaguirre C, Yeates SE, Lenz TL, Kalbe M, Milinski M. MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Mol Ecol.* 2009;18(15):3316–3329. <https://doi.org/10.1111/j.1365-294X.2009.04243.x>.
- Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A, et al. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature.* 2012;491(7426):756–760. <https://doi.org/10.1038/nature11584>.
- Endres T, Fendt M. Aversion-vs fear-inducing properties of 2, 4, 5-trimethyl-3-thiazoline, a component of fox odor, in comparison

- with those of butyric acid. *J Exp Biol.* 2009;212(15):2324–2327. <https://doi.org/10.1242/jeb.028498>.
- Feder JL, Xie X, Rull J, Velez S, Forbes A, Leung B, Dambroski H, Filchak KE, Aluja M. Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in Rhagoletis. *Proc Natl Acad Sci U S A.* 2005;102(suppl_1):6573–6580. <https://doi.org/10.1073/pnas.0502099102>.
- Fink S, Fischer MC, Excoffier L, Heckel G. Genomic scans support repetitive continental colonization events during the rapid radiation of voles (Rodentia: Microtus): the utility of AFLPs versus mitochondrial and nuclear sequence markers. *Syst Biol.* 2010;59(5):548–572. <https://doi.org/10.1093/sysbio/syq042>.
- Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Von Mering C, et al. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2012;41(D1):D808–D815. <https://doi.org/10.1093/nar/gks1094>.
- Francis F, Vanhaelen N, Haubruge E. Glutathione S-transferases in the adaptation to plant secondary metabolites in the Myzus persicae aphid. *Arch Insect Biochem Physiol.* 2005;58(3):166–174. <https://doi.org/10.1002/arch.20049>.
- Gaither MR, Bernal MA, Coleman RR, Bowen BW, Jones SA, Simison WB, Rocha LA. Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Mol Ecol.* 2015;24(7):1543–1557. <https://doi.org/10.1111/mec.13129>.
- Gamperl R. Chromosomal evolution in the genus *Clethrionomys*. *Genetica* 1982;57(3):193–197. <https://doi.org/10.1007/BF00056482>.
- Grillet M, Everaerts C, Houot B, Ritchie MG, Cobb M, Ferveur J-F. Incipient speciation in *Drosophila melanogaster* involves chemical signals. *Sci Rep.* 2012;2(1):224. <https://doi.org/10.1038/srep00224>.
- Haldane JBS. Sex ratio and unisexual sterility in hybrid animals. *J Genet.* 1922;12(2):101–109. <https://doi.org/10.1007/BF02983075>.
- Hare EE, Johnston JS. Genome size determination using flow cytometry of propidium iodide-stained nuclei. *Mol Methods Evol Genet.* 2011;772:3–12. https://doi.org/10.1007/978-1-61779-228-1_1.
- Hatfield T, Schluter D. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. *Evolution* 1999;53(3):866–873. <https://doi.org/10.2307/2640726>.
- Heckel G, Burri R, Fink S, Desmet J-F, Excoffier L. Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution.* 2005;59:2231–2242. <https://doi.org/10.1111/j.0014-3820.2005.tb00931.x>.
- Hellmann I, Prüfer K, Ji H, Zody MC, Pääbo S, Ptak SE. Why do human diversity levels vary at a megabase scale? *Genome Res.* 2005;15(9):1222–1231. <https://doi.org/10.1101/gr.3461105>.
- Hughes GM, Boston ESM, Finarelli JA, Murphy WJ, Higgins DG, Teeling EC. The birth and death of olfactory receptor gene families in mammalian niche adaptation. *Mol Biol Evol.* 2018;35(6):1390–1406. <https://doi.org/10.1093/molbev/msy028>.
- Hurst JL, Beynon RJ, Armstrong SD, Davidson AJ, Roberts SA, Gómez-Baena G, Smadja CM, Ganem G. Molecular heterogeneity in major urinary proteins of *Mus musculus* subspecies: potential candidates involved in speciation. *Sci Rep.* 2017;7(1):44992. <https://doi.org/10.1038/srep44992>.
- Jaarola M, Martinková N, Gündüz İ, Brunhoff C, Zima J, Nadachowski A, Amori G, Bulatova NS, Chondropoulos B, Fraguedakis-Tsolis S, et al. Molecular phylogeny of the speciose vole genus *Microtus* (Arvicolinae, Rodentia) inferred from mitochondrial DNA sequences. *Mol Phylogenet Evol.* 2004;33(3):647–663. <https://doi.org/10.1016/j.ympev.2004.07.015>.
- Johnson RN, O’Meally D, Chen Z, Etherington GJ, Ho SYW, Nash WJ, Grueber CE, Cheng Y, Whittington CM, Dennison S, et al. Adaptation and conservation insights from the koala genome. *Nat Genet.* 2018;50(8):1102–1111. <https://doi.org/10.1038/s41588-018-0153-5>.
- Johri P, Aquadro CF, Beaumont M, Charlesworth B, Excoffier L, Eyre-Walker A, Keightley PD, Lynch M, McVean G, Payseur BA, et al. Recommendations for improving statistical inference in population genomics. *PLoS Biol.* 2022;20(5):e3001669. <https://doi.org/10.1371/journal.pbio.3001669>.
- Klein J. Natural history of the major histocompatibility complex. New York: John Wiley; 1986.
- Klemola T, Koivula M, Korpimäki E, Norrdahl K. Small mustelid predation slows population growth of *Microtus* voles: a predator reduction experiment. *J Anim Ecol.* 1997;66(5):607–614. <https://doi.org/10.2307/5914>.
- Labutin A, Heckel G. Genome-wide support for incipient Tula hantavirus species within a single rodent host lineage. *Virus Evol.* 2024;10(1):veae002. <https://doi.org/10.1093/ve/veae002>.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature.* 2011;475(7357):493–496. <https://doi.org/10.1038/nature10231>.
- Lischer HEL, Excoffier L, Heckel G. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. *Mol Biol Evol.* 2014;31(4):817–831. <https://doi.org/10.1093/molbev/mst271>.
- Losos JB. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am Nat.* 2010;175(6):623–639. <https://doi.org/10.1086/652433>.
- Lunter G, Goodson M. Stampy: a statistical algorithm for sensitive and fast mapping of illumina sequence reads. *Genome Res.* 2011;21(6):936–939. <https://doi.org/10.1101/gr.111120.110>.
- Maibach V, Hans JB, Hvilsom C, Marques-Bonet T, Vigilant L. MHC class I diversity in chimpanzees and bonobos. *Immunogenetics* 2017;69(10):661–676. <https://doi.org/10.1007/s00251-017-0990-x>.
- Marques DA, Meier JI, Seehausen O. A combinatorial view on speciation and adaptive radiation. *Trends Ecol Evol.* 2019;34(6):531–544. <https://doi.org/10.1016/j.tree.2019.02.008>.
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD. Genome-wide evidence for speciation with gene flow in heliconius butterflies. *Genome Res.* 2013;23(11):1817–1828. <https://doi.org/10.1101/gr.159426.113>.
- Martinková N, Moravec J. Multilocus phylogeny of *arvicoline* voles (Arvicolini, Rodentia) shows small tree terrace size. *Folia Zoologica.* 2012;61(3-4):254–267. <https://doi.org/10.25225/fozo.v61.i3.a10.2012>.
- Mather N, Traves SM, Ho SYW. A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data. *Ecol Evol.* 2020;10(1):579–589. <https://doi.org/10.1002/ece3.5888>.
- Mayr E. Systematics and the origin of species, from the viewpoint of a zoologist. Harvard: Harvard University Press; 1942.
- Mazet O, Rodríguez W, Grusea S, Boitard S, Chikhi L. On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity* (Edinb). 2016;116(4):362–371. <https://doi.org/10.1038/hdy.2015.104>.
- Mazo-Vargas A, Concha C, Livraghi L, Massardo D, Wallbank RWR, Zhang L, Papador JD, Martínez-Najera D, Jiggins CD, Kronforst MR, et al. Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity. *Proc Natl Acad Sci U S A.* 2017;114(40):10701–10706. <https://doi.org/10.1073/pnas.1708149114>.

- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Meier JI, Marques DA, Wagner CE, Excoffier L, Seehausen O. Genomics of parallel ecological speciation in Lake Victoria cichlids. *Mol Biol Evol.* 2018;35(6):1489–1506. <https://doi.org/10.1093/molbev/msy051>.
- Monarca RI, Speakman JR, Mathias M. Energetics and thermal adaptation in semifossorial pine-voles *Microtus lusitanicus* and *Microtus duodecimcostatus*. *J Comp Physiol B.* 2019;189(2):309–318. <https://doi.org/10.1007/s00360-019-01205-z>.
- Nadachowska-Brzyska K, Burri R, Smeds L, Ellegren H. PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. *Mol Ecol.* 2016;25(5):1058–1072. <https://doi.org/10.1111/mec.13540>.
- Naisbit RE, Jiggins CD, Mallet J. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proc R Soc Lond B Biol Sci.* 2001;268(1478):1849–1854. <https://doi.org/10.1098/rspb.2001.1753>.
- Nei M, Li W-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A.* 1979;76(10):5269–5273. <https://doi.org/10.1073/pnas.76.10.5269>.
- Nicolaisen LE, Desai MM. Distortions in genealogies due to purifying selection and recombination. *Genetics* 2013;195(1):221–230. <https://doi.org/10.1534/genetics.113.152983>.
- Nosil P. Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. *Am Nat.* 2007;169(2):151–162. <https://doi.org/10.1086/510634>.
- Nosil P, Harmon LJ, Seehausen O. Ecological explanations for (incomplete) speciation. *Trends Ecol Evol.* 2009;24(3):145–156. <https://doi.org/10.1016/j.tree.2008.10.011>.
- O'Connor E, Strandh M, Hasselquist D, Nilsson JÅ, Westerdahl H. The evolution of highly variable immunity genes across a passerine bird radiation. *Mol Ecol.* 2016;25(4):977–989. <https://doi.org/10.1111/mec.13530>.
- Paupério J, Herman JS, Melo-Ferreira J, Jaarola M, Alves P, Searle JB. Cryptic speciation in the field vole: a multilocus approach confirms three highly divergent lineages in Eurasia. *Mol Ecol.* 2012;21(24):6015–6032. <https://doi.org/10.1111/mec.12024>.
- Pearson WR. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* 1991;11(3):635–650. [https://doi.org/10.1016/0888-7543\(91\)90071-L](https://doi.org/10.1016/0888-7543(91)90071-L).
- Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol.* 2014;31(7):1929–1936. <https://doi.org/10.1093/molbev/msu136>.
- Pierce JDJ, Ferguson B, Dewsbury DA. Conspecific preferences in prairie voles, *Microtus ochrogaster*, and meadow voles. *M. pennsylvanicus*. *Bull Psychon Soc.* 1989;27(3):267–270. <https://doi.org/10.3758/BF03334603>.
- Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Müller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, et al. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 2014;344(6190):1410–1414. <https://doi.org/10.1126/science.1253226>.
- Price TD, Bouvier MM. The evolution of F1 postzygotic incompatibilities in birds. *Evolution.* 2002;56:2083–2089. <https://doi.org/10.1111/j.0014-3820.2002.tb00133.x>.
- Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW, et al. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* 2016;26(3):342–350. <https://doi.org/10.1101/gr.193474.115>.
- Quina A, Bastos-Silveira C, Miñarro M, Ventura J, Jiménez R, Paulo O, da Luz Mathias M. P53 gene discriminates two ecologically divergent sister species of pine voles. *Heredity (Edinb).* 2015;115(5):444–451. <https://doi.org/10.1038/hdy.2015.44>.
- Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun.* 2013;4(1):1958. <https://doi.org/10.1038/ncomms2958>.
- Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* 2000;16(6):276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2).
- Rosário IT, Cardoso PE, Mathias M. Is habitat selection by the Cabrera vole (*Microtus cabreræ*) related to food preferences? *Mamm Biol.* 2008;73(6):423–429. <https://doi.org/10.1016/j.mambio.2008.05.001>.
- Rundle HD, Nosil P. Ecological speciation. *Ecol Lett.* 2005;8(3):336–352. <https://doi.org/10.1111/j.1461-0248.2004.00715.x>.
- Saito H, Nishizumi H, Suzuki S, Matsumoto H, Ieki N, Abe T, Kiyonari H, Morita M, Yokota H, Hirayama N, et al. Immobility responses are induced by photoactivation of single glomerular species responsive to fox odour TMT. *Nat Commun.* 2017;8(1):16011. <https://doi.org/10.1038/ncomms16011>.
- Saxenhofer M, Labutin A, White TA, Heckel G. Host genetic factors associated with the range limit of a European hantavirus. *Mol Ecol.* 2022;31(1):252–265. <https://doi.org/10.1111/mec.16211>.
- Saxenhofer M, Schmidt S, Ulrich RG, Heckel G. Secondary contact between diverged host lineages entails ecological speciation in a European hantavirus. *PLoS Biol.* 2019;17(2):e3000142. <https://doi.org/10.1371/journal.pbio.3000142>.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015;31(19):3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
- Smadja C, Butlin R. On the scent of speciation: the chemosensory system and its role in premating isolation. *Heredity (Edinb).* 2009;102(1):77–97. <https://doi.org/10.1038/hdy.2008.55>.
- Smadja C, Catalan J, Ganem G. Strong premating divergence in a unimodal hybrid zone between two subspecies of the house mouse. *J Evol Biol.* 2004;17(1):165–176. <https://doi.org/10.1046/j.1420-9101.2003.00647.x>.
- Smadja CM, Loire E, Caminade P, Thoma M, Latour Y, Roux C, Thoss M, Penn DJ, Ganem G, Boursot P. Seeking signatures of reinforcement at the genetic level: a hitchhiking mapping and candidate gene approach in the house mouse. *Mol Ecol.* 2015;24(16):4222–4237. <https://doi.org/10.1111/mec.13301>.
- Smit AFA, Hubble R. RepeatModeler Open-1.0. <http://www.repeatmasker.org>. 2008.
- Smit AFA, Hubble R, Green P. 2015. RepeatMasker Open-4.0. p. 2013–2015. <http://www.repeatmasker.org>.
- Soares JMA. Fitness dos híbridos entre duas espécies de ratos fossadores, *Microtus lusitanicus* (Gerbe, 1879) e *M. Duodecimcostatus* (de Selys-Longchamps, 1839): uma abordagem multidisciplinar. PhD-thesis. University of Lisbon; 2013.
- Soni V, Johri P, Jensen JD. Evaluating power to detect recurrent selective sweeps under increasingly realistic evolutionary null models. *Evolution* 2023;77(10):2113–2127. <https://doi.org/10.1093/evolut/qpad120>.
- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30(9):1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.

- Stanke M, Waack S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 2003;19(suppl_2):ii215–ii225. <https://doi.org/10.1093/bioinformatics/btg1080>.
- Stelkens RB, Young KA, Seehausen O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* 2010;64(3):617–633. <https://doi.org/10.1111/j.1558-5646.2009.00849.x>.
- Sutouh Y, Mohamed RH, Kasahara M. Origin and evolution of dendritic epidermal T cells. *Front Immunol*. 2018;9:1059. <https://doi.org/10.3389/fimmu.2018.01059>.
- Sutter A, Beysard M, Heckel G. Sex-specific clines support incipient speciation in a common European mammal. *Heredity (Edinb)*. 2013;110(4):398–404. <https://doi.org/10.1038/hdy.2012.124>.
- Svedin N, Wiley C, Veen T, Gustafsson L, Qvarnström A. Natural and sexual selection against hybrid flycatchers. *Proc R Soc Lond B Biol Sci*. 2008;275(1635):735–744. <https://doi.org/10.1098/rspb.2007.0967>.
- Teshima KM, Coop G, Przeworski M. How reliable are empirical genomic scans for selective sweeps? *Genome Res*. 2006;16(6):702–712. <https://doi.org/10.1101/gr.5105206>.
- Thornton KR, Jensen JD. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* 2007;175(2):737–750. <https://doi.org/10.1534/genetics.106.064642>.
- Torgasheva AA, Borodin PM. Cytological basis of sterility in male and female hybrids between sibling species of grey voles *Microtus arvalis* and *M. levis*. *Sci Rep*. 2016;6(1):36564. <https://doi.org/10.1038/srep36564>.
- Tran VDT, De Coi N, Feuermann M, Schmid-Siegert E, Bağcı E-T, Mignon B, Waridel P, Peter C, Pradervand S, Pagni M, et al. RNA sequencing-based genome reannotation of the dermatophyte *Arthroderma benhamiae* and characterization of its secretome and whole gene expression profile during infection. *MSystems* 2016;1(4):e00036–e00016. <https://doi.org/10.1128/mSystems.00036-16>.
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562–578. <https://doi.org/10.1038/nprot.2012.016>.
- Turelli M, Moyle LC. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics* 2007;176(2):1059–1088. <https://doi.org/10.1534/genetics.106.065979>.
- Wagner CE, Harmon LJ, Seehausen O. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 2012;487(7407):366–369. <https://doi.org/10.1038/nature11144>.
- Wakeley J. The coalescent in an island model of population subdivision with variation among demes. *Theor Popul Biol*. 2001;59(2):133–144. <https://doi.org/10.1006/tpbi.2000.1495>.
- Wang X, Heckel G. Genome-wide relaxation of selection and the evolution of the island syndrome in Orkney voles. *Genome Res*. 2024;34(6):851–862. <https://doi.org/10.1101/gr.278487.123>.
- Wang X, Peischl S, Heckel G. Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal. *Curr Biol*. 2023;33(10):2051–2062.e2054. <https://doi.org/10.1016/j.cub.2023.04.042>.
- Wolf JB, Ellegren H. Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet*. 2017;18(2):87. <https://doi.org/10.1038/nrg.2016.133>.
- Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, et al. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 2010;329(5987):75–78. <https://doi.org/10.1126/science.1190371>.
- Yukilevich R, Harvey T, Nguyen S, Kehlbeck J, Park A. The search for causal traits of speciation: divergent female mate preferences target male courtship song, not pheromones, in *Drosophila* *Athabasca* species complex. *Evolution* 2016;70(3):526–542. <https://doi.org/10.1111/evo.12870>.

Associate editor: Susanne Pfeifer