



OPEN

Whole-exome sequencing and genome-wide evolutionary analyses identify novel candidate genes associated with infrared perception in pit vipers

Na Tu, Dan Liang & Peng Zhang

Pit vipers possess a unique thermal sensory system consisting of facial pits that allow them to detect minute temperature fluctuations within their environments. Biologists have long attempted to elucidate the genetic basis underlying the infrared perception of pit vipers. Early studies have shown that the *TRPA1* gene is the thermal sensor associated with infrared detection in pit vipers. However, whether genes other than *TRPA1* are also involved in the infrared perception of pit vipers remains unknown. Here, we sequenced the whole exomes of ten snake species and performed genome-wide evolutionary analyses to search for novel candidate genes that might be involved in the infrared perception of pit vipers. We applied both branch-length-comparison and selection-pressure-alteration analyses to identify genes that specifically underwent accelerated evolution in the ancestral lineage of pit vipers. A total of 47 genes were identified. These genes were significantly enriched in the ion transmembrane transporter, stabilization of membrane potential, and temperature gating activity functional categories. The expression levels of these candidate genes in relevant nerve tissues (trigeminal ganglion, dorsal root ganglion, midbrain, and cerebrum) were also investigated in this study. We further chose one of our candidate genes, the potassium channel gene *KCNK4*, as an example to discuss its possible role in the infrared perception of pit vipers. Our study provides the first genome-wide survey of infrared perception-related genes in pit vipers via comparative evolutionary analyses and reveals valuable candidate genes for future functional studies.

Infrared perception is a notable evolutionary adaptation of snakes that helps snakes respond to the external environment during their nocturnal activities and prey on warm-blooded animals^{1–3}. Pit vipers are the best at sensing and utilizing infrared signals among all snakes^{4–6}. They have the most sensitive infrared sensory system, which can detect minute temperature fluctuations as low as 0.001 °C in their immediate surroundings⁷. This peculiar infrared perception ability enables pit vipers to detect and locate warm-blooded animals with high precision in the dark, which makes them one of the most successful predators in nature^{8–10}.

Biologists have long sought to investigate the biological basis of infrared perception in pit vipers. Previous anatomical and neurophysiological studies have accumulated substantial data explaining the detection and transduction mechanism of pit viper infrared perception^{6,11–17}. We now know that the infrared perception of pit vipers is initiated by the heat stimulation of a thin sensory membrane embedded in the facial loreal pit, a highly specialized organ located between the eye and nostril^{11,12}. The sensory membrane is innervated by afferent fibres of trigeminal nerve branches that transduce infrared signals to a particular nucleus in the hindbrain^{13–15}. The efferents of the nucleus then project to the optic tectum, where infrared information is integrated with visual information to construct a thermal image (also known as “infrared vision”)^{16,17}. This complex infrared sensory system must involve many genes at the genetic level. However, the genetic basis of the system that is involved in the infrared perception of pit vipers is far from defined and requires more attention. Recently, a study by Gracheva et al.¹⁸ showed that a particular gene, *TRPA1*, is highly expressed in the trigeminal ganglion (TG) of pit vipers.

State Key Laboratory of Biocontrol, College of Ecology and Evolution, School of Life Sciences, Higher Education Mega Center, Sun Yat-Sen University, Guangzhou 510006, China. email: zhangp35@mail.sysu.edu.cn

This gene encodes a calcium ion channel that can respond to temperature changes caused by infrared stimulation. Thus far, *TRPA1* is the only gene that has been shown to be involved in the infrared perception of pit viper snakes.

In pit vipers (and other infrared-sensitive snakes such as pythons and some boas⁶), the expression level of *TRPA1* in the TG is much higher than that in the dorsal root ganglion (DRG), which is not observed in infrared-insensitive snakes¹⁸. This difference in expression was the cornerstone of the transcriptome profiling analysis that led to the discovery of *TRPA1*. This strategy is able to reveal direct thermal sensor proteins that are involved in the infrared perception of snakes. However, it may not be able to identify genes involved in pit organ evolution, signal amplification, or information integration (if any) during infrared perception because such genes may not exhibit considerable changes in expression in different nerve tissues.

Taxonomically, pit vipers (Crotalinae) belong to the family Viperidae. There are three major clades of Viperidae: true vipers (Viperinae), pit vipers (Crotalinae), and Fea's vipers (Azemiopinae), among which Viperinae diverged from the others first, followed by Azemiopinae and Crotalinae^{19,20}. Notably, true vipers and Fea's vipers do not have pit organs, and early animal behavior studies have found that these snakes are unable to sense infrared radiation^{21,22}. These two viper groups are rather similar to pit vipers in many aspects, such as their morphology, venom, and lifestyle, but do not possess infrared perception ability²³. Therefore, the infrared perception ability of pit vipers must have evolved after they diverged from the other two types of vipers.

Phenotype-associated genes often exhibit a particular evolutionary pattern in taxa with that phenotype compared to their sister taxa without the phenotype²⁴. For example, the "language gene" *FOXP2* underwent accelerated evolution in the human lineage after its separation from the lineage leading to chimpanzees and bonobos, our closest relatives^{25–27}; the echolocation gene *Prestin* shows convergent evolution in echolocating bats and whales^{28–30}. Such comparative evolutionary analyses are an effective strategy for searching for phenotype-associated genes. Notably, our earlier work showed that the thermoreceptor gene *TRPA1* underwent strong accelerated evolution in pit vipers but not in true vipers³¹. In contrast, *TRPV1*, another TRP gene that is not associated with infrared perception, shows no difference in the rate of evolution between pit vipers and true vipers. Moreover, *TRPA1* contains many amino acid (AA) residues that are unique in pit vipers but evolutionarily conserved in other snakes³¹. We hypothesize that if there are other genes involved in the infrared perception of pit vipers, these genes may exhibit a similar evolutionary pattern to *TRPA1*. This particular evolutionary pattern could be used as a screening criterion to search for potential genes associated with infrared perception in pit vipers, which may compensate for the limitations of transcriptome profiling analysis.

A number of snake genomes have been sequenced and released in public databases^{32–37}, including the genomes of members of both Viperinae and Crotalinae. However, representative genomes from Azemiopinae are lacking, hindering genome-wide comparative studies for the identification of infrared perception-related genes. In this study, we performed whole-exome capture sequencing to supplement the genomic data for the three subfamilies of Viperidae and other related snake species. Based on these new genomic data, we conducted a genome-wide survey to search for novel candidate genes related to the infrared perception of pit vipers. Our comparative evolutionary analyses identified 47 candidate genes with evolutionary patterns similar to that of *TRPA1*, which may be associated with infrared perception in pit vipers. Most of these genes are associated with ion transmembrane transporter activity. We hope that the genome-wide identification and evolutionary characterization of novel candidate genes will contribute to the further elucidation of the genetic basis of the infrared perception of pit vipers.

Results and discussion

Coding sequence (CDS) data. In this study, we sequenced the whole coding sequences of a total of 10 snake species (see Supplementary Table S1 for sample details) by sequence capture. The sequence capture experiment was performed by using a custom-designed capture probe set (Agilent Custom SureSelect Kit) to selectively target the coding regions of ~20,000 snake protein-coding genes. High-throughput sequencing was performed on an Illumina HiSeq X-10 sequencer in 150-bp paired-end mode, and the numbers of read pairs among the samples ranged from 27 to 56 million (40.25 million on average). Approximately 23% of the reads could be mapped to the 21,331 reference CDSs of *Deinagkistrodon acutus*³², and the on-target percentage across the ten snake samples varied from 9.8 to 31.3%. Based on these data, we were able to reconstruct over 20,000 CDSs across all the samples (average = 21,197) with an average CDS data completeness of 92.24% (Supplementary Table S1). We compared the obtained CDSs of our *Ophiophagus hannah* sample with the sequences of the previously published king cobra genome³⁸ and found 99% sequence similarity, suggesting that the sequence quality of our CDS data is high. By combining these newly generated CDS data with the 9 published snake genomes (Supplementary Table S2), we produced a dataset consisting of 21,331 orthologous gene groups (OGGs) for the 19 snake species, including eight pit vipers, one Fea's viper, two true vipers, and representative species from five other snake families.

General pattern of gene evolution in viper snakes. To better understand the evolutionary history of the viper snakes included in this study and provide an accurate phylogenetic framework for comparative analysis, we first constructed a genome-scale phylogenetic tree for the 19 snake species based on the concatenated protein sequences of all 21,331 OGGs. The resulting tree (Fig. 1a) was highly resolved, with all nodes exhibiting 100% bootstrap support, and was consistent with a recent snake phylogeny^{19,20}. In this tree, Viperidae is monophyletic; Azemiopinae (Fea's vipers) is the sister group of Crotalinae (pit vipers), and Viperinae (true vipers) is the sister group of both of the first two groups. Notably, within Viperidae, the branch lengths from the terminal tips to the root node of Viperidae are rather uniform across the three subfamilies, suggesting that most of the protein-coding genes evolved at a similar pace among the three types of viper snakes. In contrast, in the gene tree of the infrared perception-related *TRPA1* gene (Fig. 1b), the branch lengths of Crotalinae are considerably longer than those of the other two viper subfamilies.

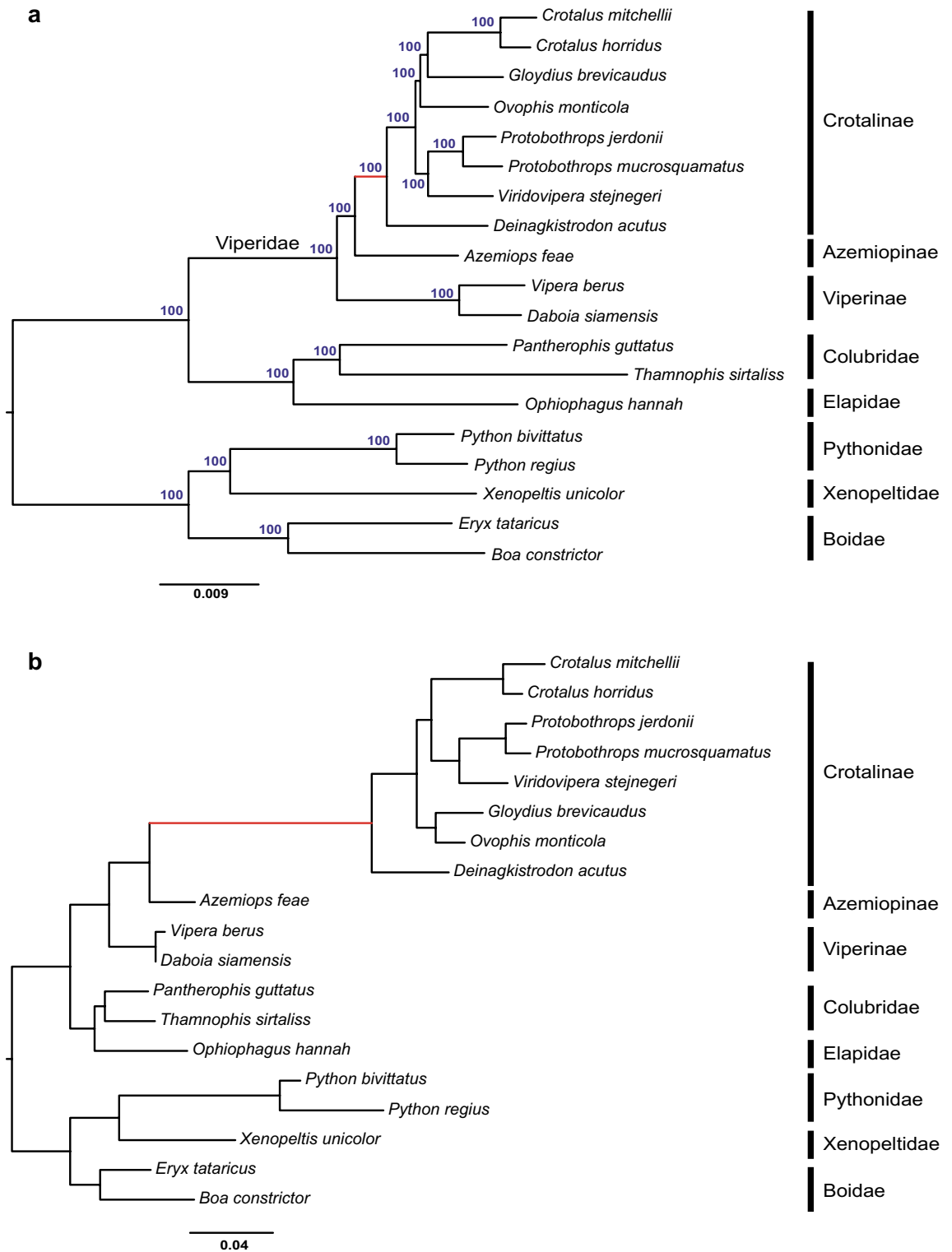


Figure 1. The genome phylogenetic tree (a) and the *TRPA1* gene tree (b). Numbers on the branches are ML bootstrap support values. The ancestral branch of pit vipers are marked in red.

Identification of candidate genes associated with infrared perception in pit vipers. In the *TRPA1* gene tree, the ancestral branch of pit vipers is considerably longer (more than three times) than the sister branch connecting to true vipers, indicating that pit vipers have evolved considerably faster than all other snakes (exemplar tree; Fig. 2). We thus first utilized this special branch length pattern to search for potential infrared perception-related genes among the 21,331 OGGs.

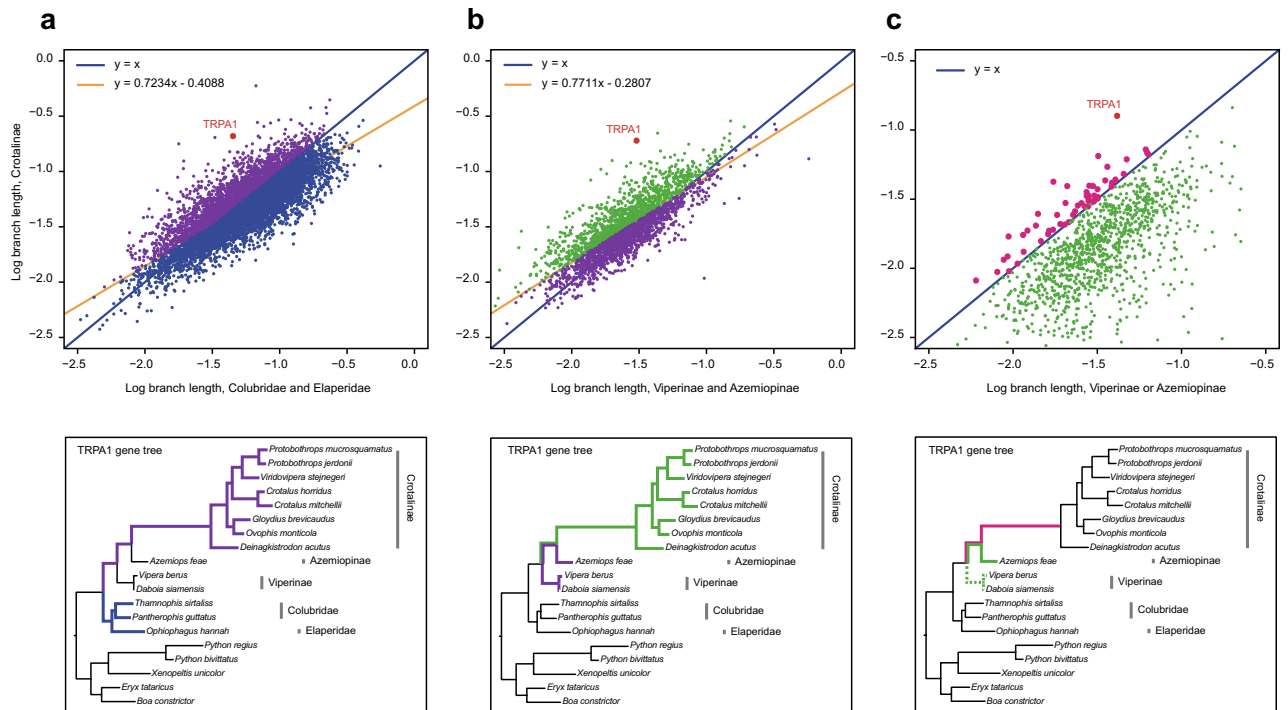


Figure 2. Genome-wide survey of candidate genes for infrared perception in pit vipers based on branch length analysis. The analysis includes three comparisons performed to search for genes that exhibit similar evolutionary patterns to *TRPA1*. The *TRPA1* gene tree is provided as an example below every search step. (a) The average root-to-tip branch length of pit vipers (purple branches) compared to the average root-to-tip branch length of the Colubridae + Elapidae clade (blue branches). Each dot represents one gene. Purple dots are retained for the next step of filtering. (b) The average root-to-tip branch length of pit vipers (green branches) compared to that of other vipers (purple branches). Green dots are retained for the next step of filtering. (c) The ancestral branch length of pit vipers (magenta branches) compared to the longest root-to-tip branch length of other viper snakes (green branches). The longest branch is highlighted by the solid green line, whereas the dashed green lines represent shorter branches. Magenta dots represent the final retained genes.

In this analysis, individual gene trees were reconstructed for all 21,331 OGGs by maximum likelihood inference using the obtained genome tree as a topology constraint. We used a three-step filtering strategy to gradually narrow our candidate gene pool from these gene trees. First, we wanted to screen out genes that evolved more rapidly in Crotalinae (pit vipers) than in Colubridae and Elapidae (outgroup snakes). To do so, we considered the average branch length from the common ancestral node to all termini of a given snake group as the average evolution rate of the group. For example, starting from the common ancestral node of Crotalinae + Colubridae + Elapidae, the average branch length to all Crotalinae termini (Fig. 2a; purple branches in the phylogram) and the average branch length to all Colubridae + Elapidae termini (Fig. 2a; blue branches in the phylogram) were regarded as the average evolution rates of Crotalinae and the outgroup, respectively. For both snake groups, the average branch lengths were calculated and plotted. The genes located above the regression line and the 1:1 line (evolving faster in pit vipers than in outgroup snakes) were retained (Fig. 2a). Second, the retained genes were further filtered by comparing the average branch length of Crotalinae with that of Viperinae and Azemiopinae (infrared-insensitive viper snakes) (Fig. 2b). Genes located below the 1:1 geometric and regression lines were filtered out. By doing so, we aimed to ensure that the selected candidate genes have evolved faster in pit vipers than in the other two types of viper snakes. Third, we plotted the ancestral branch length of pit vipers and the longest branch length leading to the Viperinae or Azemiopinae terminal and retained genes located above the 1:1 line (Fig. 2c). This step was performed to identify genes that experienced accelerated evolution on the ancestral branch of pit vipers. The three-step filtering process led to the identification of a total of 51 candidate genes. Finally, we reconstructed ML trees for these 51 genes without topology constraints and repeated the above three filtering steps. Among the 51 candidate genes, 38 genes passed the filtration procedure (Supplementary Table S3).

We manually inspected the gene tree of these 38 candidate genes to check for pit-viper-specific signatures of accelerated evolution (Supplementary Fig. S1). All these gene trees exhibit similar evolutionary patterns to *TRPA1*, although the patterns are somewhat less conspicuous than that of *TRPA1*. For example, the *KCNK4* gene displays a relatively longer length along the common ancestral branch of pit vipers, and the total evolutionary branch length is significantly longer than those of all other snake lineages; the *KCNH8* gene displays a longer branch length not only in pit vipers but also in pythons (another group of snakes with infrared perception ability), which may imply a close association with their common characteristic of infrared perception ability; a hypothetical gene with the ID XB-GENE-5885309 has evolved much faster in pit vipers than in adjacent vipers. These

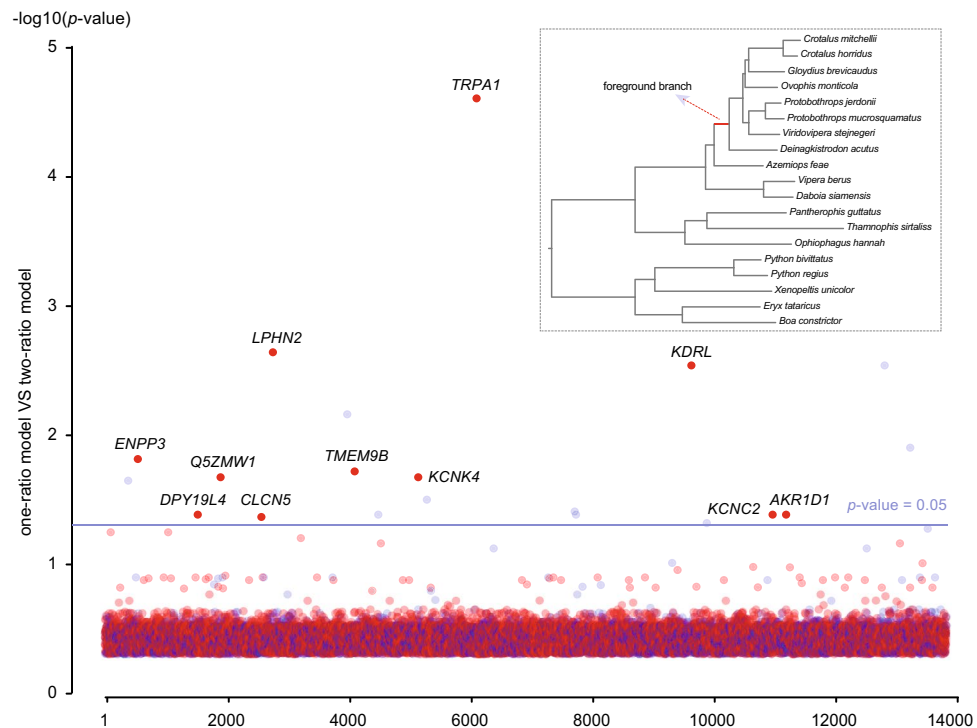


Figure 3. Genome-wide survey of candidate genes for infrared perception in pit vipers based on selection pressure analysis. The one-ratio model is compared with the two-ratio model, and their relative fit is evaluated using likelihood ratio tests with statistical correction. The ancestral branch of the pit viper lineage is set as the foreground branch. Genes with $\omega_1 > \omega_0$ are indicated in red, whereas genes with $\omega_1 < \omega_0$ are indicated in light blue. Each dot represents an individual gene, and the x-axis represents the numerical order of the genes.

shared evolutionary patterns similar to *TRPA1* suggest that other genes might also participate in the specialized trait of pit vipers, the infrared perception.

In addition to the branch length analysis, we performed a genome-wide scan based on selection pressure comparison. As an infrared perception-related gene of pit vipers, *TRPA1* experienced significantly different selection pressure, showing an elevated dN/dS value (the ratio of nonsynonymous to synonymous substitutions) in the ancestral branch of pit vipers³¹. Thus, we compared the one-ratio model and the two-ratio model for all 21,331 OGGs by PAML³⁹. The one-ratio model assumes that all branches in the phylogeny have the same ω value (dN/dS), while the two-ratio model assumes two different ω values within the phylogeny: one for the ancestral branch of pit vipers (ω_1) and one for the remaining branches of the tree (ω_0). If the two-ratio model is significantly better than the one-ratio model for a gene, the gene experienced significantly different selection pressure in the ancestral branch of pit vipers. The selection pressure analysis identified a total of 11 genes with significantly elevated ω values in the ancestral branch of pit vipers compared to the background branches ($\omega_1 > \omega_0$, with FDR-adjusted P-value < 0.05; Fig. 3 and Supplementary Tables S4 and S5).

Combining the 11 candidate genes from the selection analysis and the 38 candidate genes from the branch length analysis, we identified a total of 47 candidate genes that may be related to the infrared perception of pit vipers. These genes are *TRPA1*, *KCNK4*, *WAPAL*, *NOC4L*, *Q5ZLE7*, *ENPP3*, *KCNH8*, *LPHN2*, *TMEM9B*, *XB-GENE-5885309*, *CLCN5*, *DPY19L4*, *GK5*, *KDRL*, *ADAMTS2*, *AKR1D1*, *HRH2*, *C1QL2*, *CAST*, *HABP4*, *LHX3*, *SLC6A16*, *ZNF703*, *ZNT6*, *CERS6*, *COQ5*, *F1NPK8*, *FAM58A*, *KCNC2*, *KCNK17*, *METTL20*, *Q5ZMW1*, *SETD7*, *SLC45A1*, *VWA7*, *WVOX*, *AGTR1*, *CHMP6*, *CRYAB*, *CYBB*, *DECR2*, *LRRC3*, *MYOT*, *NDUF8*, *SKA1*, *SLC5A1*, and *ZGC_136493*. Among these genes, only two (*TRPA1* and *KCNK4*) were identified simultaneously by both filtered schemes. Detailed information on all candidate genes is given in Supplementary Table S6.

Characteristics of the candidate genes. To further explore the possible associations between the identified genes and the infrared perception function in pit vipers, we retrieved Gene Ontology (GO) terms for the 47 candidate genes under three major categories: biological process (BP), cellular component (CC), and molecular function (MF). Additionally, we performed GO enrichment analysis to test whether our candidate genes were significantly enriched in specific GO terms. The GO enrichment analysis showed that five BP GO terms, two CC GO terms, and four MF GO terms were significantly enriched (Table 1). A GO term of the CC category (integral component of membrane) included the most genes (17 of the candidate genes). Remarkably, all enriched GO terms of the MF category were related to ion channel genes, including the *TRPA1*, *KCNK4*, *KCNH8*, *KCNK17*, *KCNC2*, *CYBB*, and *CRYAB* genes. Ion channel proteins are crucial in the thermosensation of vertebrates⁴⁰. Because our screening scheme targeted infrared perception, these ion channel genes are unlikely to have been

Categories	Description	Gene number	Corrected <i>P</i> value	Enrichment fold	Genes
Biological progress					
GO:0071805	Potassium ion transmembrane transport	4	1.10E-04	15.9	KCNK17, KCNH8, KCNK4, KCNC2
GO:0034765	Regulation of ion transmembrane transport	4	8.10E-05	17.3	KCNK17, CYBB, KCNH8, KCNK4
GO:0030322	Stabilization of membrane potential	2	5.00E-04	60	KCNK17, KCNK4
GO:0042542	Response to hydrogen peroxide	2	5.00E-03	18.8	CRYAB, TRPA1
GO:0014070	Response to organic cyclic compound	2	4.70E-03	19.6	KCNC2, TRPA1
Cellular component					
GO:0005887	Integral component of plasma membrane	12	1.40E-05	4.2	AGTR1, CLCN5, CYBB, ENPP3, HRH2, KCNK17, KCNK4, KCNC2, KCNH8, SLC5A1, SLC6A16, TRPA1
GO:0016021	Integral component of membrane	17	1.70E-02	1.6	NDUFB8, TMEM9B, AGTR1, CERS6, CLCN5, CYBB, DPY19L4, ENPP3, LRR3, NOC4L, KCNK17, KCNK4, KCNC2, SLC45A1, SLC5A1, SLC6A16, TRPA1
Molecular function					
GO:0005244	Voltage-gated ion channel activity	3	4.20E-05	44	KCNK17, CYBB, KCNK4
GO:0097604	Temperature-gated cation channel activity	2	4.40E-06	468.9	TRPA1, KCNK4
GO:0022841	Potassium ion leak channel activity	2	5.20E-04	58.6	KCNK17, KCNK4
GO:0005267	Potassium channel activity	2	2.80E-03	25.3	KCNK17, KCNK4

Table 1. Over-represented GO categories among the 47 candidate genes.

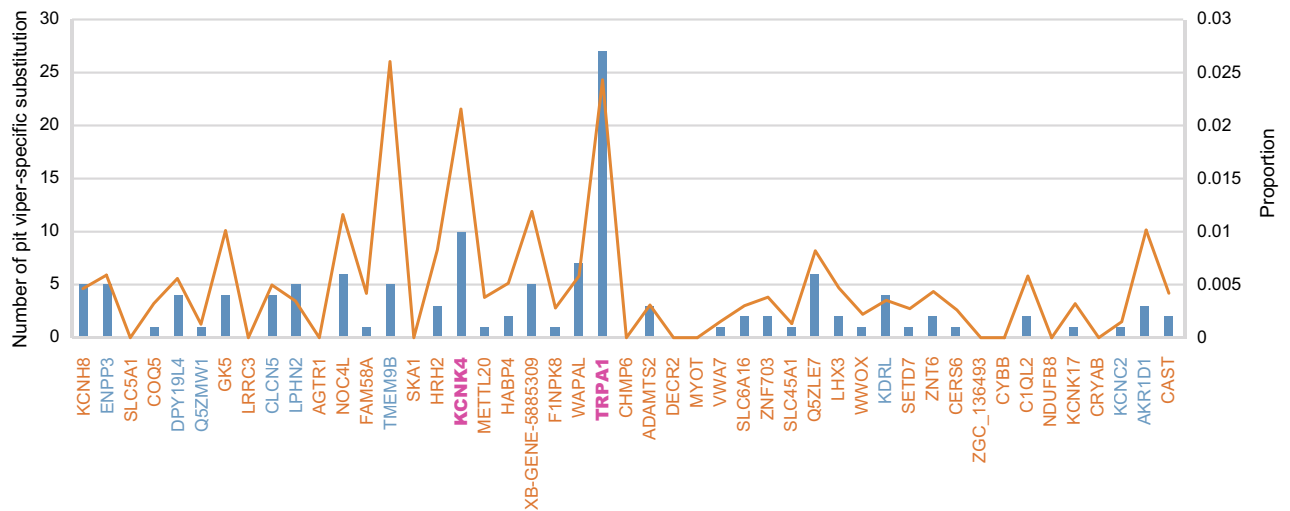


Figure 4. The number and proportion of pit viper-specific amino acid substitutions for the 47 candidate genes. The blue bars are numbers, and the orange lines are proportions. The genes identified by branch length analysis are orange, the genes identified by selection analysis are blue, and the two genes (*TRPA1* and *KCNK4*) identified by both analyses are magenta.

randomly identified. It is worth mentioning that we found that the two genes (*TRPA1* and *KCNK4*) that were identified by both the selection pressure and branch length analyses were associated with the GO term ‘temperature-gated cation channel activity (GO: 0097604)’. This GO term was significantly enriched (468.9-fold), with the lowest corrected *P*-value of 4.4e-6.

Residue substitutions at evolutionarily conserved positions are often signs of the functional alteration of proteins. Therefore, we calculated the number of pit viper-specific amino acid (AA) substitutions for each candidate gene and used a branch-specific site model to test whether these substitution sites are under positive selection. A pit viper-specific substitution is defined as an amino acid site that is in one state for all other snake groups but is substituted to another state in pit vipers. Thirty-six of the 47 candidate genes possess at least one pit viper-specific substitution; specifically, the *TRPA1*, *KCNK4*, *WAPAL*, *NOC4L*, *Q5ZLE7*, *ENPP3*, *KCNH8*, *TMEM9B*, *XB-GENE-5885309*, *CLCN5*, *DPY19*, *GK5*, *KDRL*, and *ADAMTS2* genes present more than five pit viper-specific AA substitutions (Fig. 4 and Supplementary Table S6). Most of these substitution sites are also under positive selection in the ancestral branch of pit vipers (Supplementary Table S6). The *TRPA1* gene possesses the highest number of pit viper-specific AA substitutions (27 in total) in its protein sequence. Among the remaining candidate genes, the *KCNK4* gene possesses the highest number of pit viper-specific AA substitutions (10 in total), and the proportion is comparable to that for *TRPA1* (Fig. 4). When considered according to the proportion of pit viper-specific AA substitutions, *TMEM9B* is the most prominent gene among all candidate genes. This gene

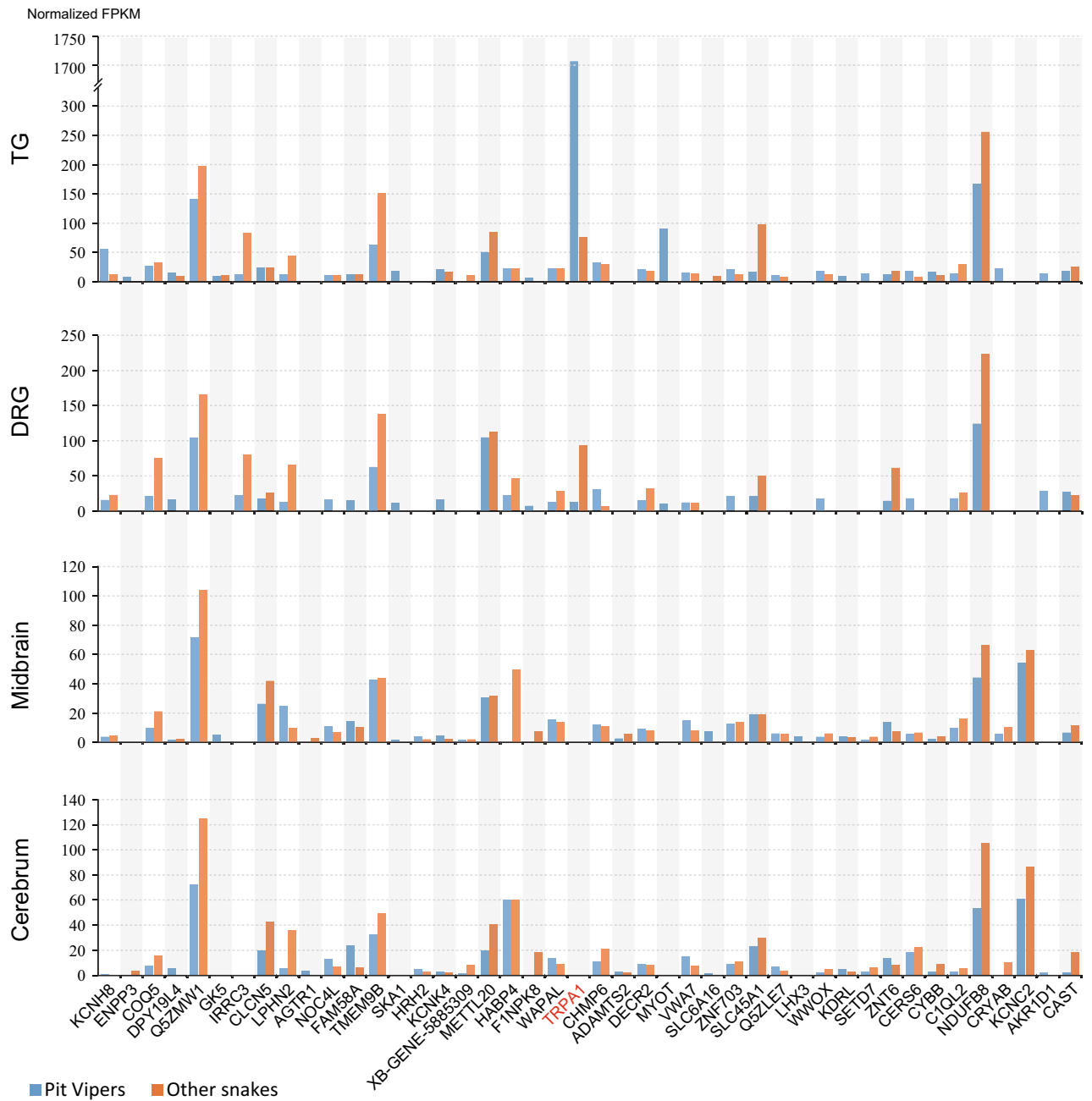


Figure 5. Expression levels of candidate genes in four nerve tissues of pit vipers and other infrared-insensitive snakes: trigeminal ganglion (TG), dorsal root ganglion (DRG), cerebrum, and midbrain.

is related to an obsolete signal transducer activity (GeneCards, <https://www.genecards.org>) and exhibits a functional interaction with *CLCN5*, a gene that is also present in our candidate gene set. In summary, the functional properties of many of our candidate genes suggest that they may play a role in the infrared perception of pit vipers, but their functional mechanisms need to be explored in more detail.

Expression levels of candidate genes in cerebrum, midbrain, TG and DRG nerve tissues. Our comparative evolutionary analyses identified a total of 47 candidate genes. The investigation of the expression levels of these candidate genes in relevant nerve tissues is necessary to elucidate their functional roles in infrared perception in pit vipers. Therefore, we performed transcriptome profiling to investigate the expression levels of the candidate genes in the cerebrum, midbrain, TG, and DRG nerve tissues of both pit vipers and infrared-insensitive snakes (Fig. 5). In pit vipers, the TG contains sensory neurons that directly connect to the sensory membrane embedded in the facial pit. The DRG, midbrain, and cerebrum were used as comparative references to the TG. Expression levels were measured by the trimmed mean of M-values (TMM) normalized fragments per kilobase of transcript per million mapped reads (FPKM) values to compare the expression levels of different genes across different samples.

Among the 47 candidate genes, three genes (*KCNK17*, *SLC5A1*, and *ZGC_136493*) were not expressed in any of the four nerve tissues that we investigated. Among the remaining 44 genes, all but four genes (*AGTR1*, *ADAMTS2*, *LHX3*, and *KCNC2*) were expressed in the TG of pit vipers. When pit vipers were compared with infrared-insensitive snakes, nine genes were found to exhibit upregulated expression in the pit viper TG. They are *TRPA1*, *KCNH8*, *MYOT*, *KCNK4*, *ZNF703*, *KDRL*, *SETD7*, *CYBB*, and *CRYAB* (Fig. 5). For example, the expression level of *TRPA1* in the pit viper TG (FPKM 1707.7) is considerably higher than in the other three nerve tissues (DRG: 12.7; midbrain: 0; cerebrum: 0). The expression levels of *KCNH8* and *MYOT* in the pit viper TG and the other three nerve tissues are TG: 56.5; DRG: 15.6; midbrain: 3.8; cerebrum: 1.1 and TG: 90.5; DRG: 10.1; midbrain: 0; cerebrum: 0, respectively. The expression level of *KCNK4* in the TG is also higher than the levels in the other three nerve tissues, but the expression differences are small (TG: 21.2; DRG: 16.3; midbrain: 4.7; cerebrum: 2.7). Detailed quantitative results are provided in Supplementary Table S7.

Among the eight upregulated genes (excluding *TRPA1*), the molecular functions of five genes (*ZNF703*, *KDRL*, *SETD7*, *CYBB*, and *CRYAB*) show no direct association with infrared perception based on the current knowledge of human orthologous proteins. Among the remaining three genes, *KCNH8* is a voltage-gated potassium channel gene involved in regulating neurotransmitter release and neuronal excitability⁴¹; *MYOT* encodes a cytoskeletal protein that plays an essential role in stabilizing thin filaments during muscle contraction⁴², and *KCNK4* functions as an outwardly rectifying channel and may be involved in regulating the input threshold in sensory neurons⁴³. All of these functions are associated with infrared perception to some extent. The relatively higher expression of these genes in the trigeminal ganglion of pit vipers implied that they might function in the infrared perception of pit vipers.

Possible functional mechanism with *KCNK4* as an example. Through genome-wide evolutionary analyses, we identified a number of genes that may function during the infrared perception of pit vipers. However, whether these genes are truly functional is still unclear. We thus attempted to obtain some clues regarding the function of these genes via protein structure simulation and functional inference. *KCNK4* was a repeatedly identified candidate gene in all of our evolutionary analyses. This gene has experienced accelerated evolution in the ancestral branch of pit vipers and accumulated large numbers of pit viper-specific amino acid mutations (Supplementary Fig. S2), and it exhibits upregulated expression in pit viper TG. Although the upregulated expression pattern of this gene in the pit viper TG was not significant (Fig. 5), the unique amino acid mutations distributed in critical regions of the protein may confer a functional advantage in the infrared perception of pit vipers. We, therefore, chose *KCNK4* as an example to perform protein structure simulation to further explore its possible functional mechanism.

The protein structure simulation showed that the complete *KCNK4* channel is a homodimer in which each subunit comprises an extracellular cap that controls channel egress, four transmembrane helices (M1–M4) surrounding the central cavity, and two pore-helices (PH1 and PH2) forming the selectivity filter (Fig. 6a, b, d)⁴⁴. Among the ten pit viper-specific amino acid mutations (Fig. 6c), three (R38W, A43T, I47V) are located in M1, and two (L277M, I283V) are located in M4. Earlier studies reported that amino acid mutations in the transmembrane helices of K_{2p} channels would affect channel gating, which may blunt the channel responses to various types of gating commands^{45,46}. Therefore, we hypothesized that the five amino acid mutations in the transmembrane helices of pit viper *KCNK4* might have a similar effect. Pore helices are vital components that play an essential role in maintaining the pore conformation of the *KCNK4* channel^{44,47}. Mutations in the pore helices of various K_{2p} channels, such as *KCNK1*, *KCNK2*, *KCNK4*, and *KCNK9*, have been reported to dramatically affect the gating of the channels^{45,48,49}. Pit viper *KCNK4* harbors one mutation (S126N) located in pore helix 1 (PH1). Protein structure simulation showed that the side chain of this amino acid residue points towards the central pore. In terms of space, the asparagine (N) side chain is bulkier than the serine (S) side chain, causing increased steric hindrance (Fig. 6e), which is also likely to have an impact on channel gating. Three additional amino acid mutations (E94A, D109N, S259R) are located in the extracellular cap helix, the upstream linker of PH1, and the downstream linker of PH2. Protein structure simulation showed that these three residues are distributed around the channel egress (Fig. 6f). Interestingly, in pit viper *KCNK4*, two negatively charged residues, Glu and Asp, are replaced with two noncharged residues, Ala and Asn (E94A, D109N), and the noncharged residue Ser is replaced by the positively charged residue Arg (S259R). Such replacements alter the charge environment of the channel egress from a negative state, which attracts potassium ions, to a positive state, which repels potassium ions, potentially also affecting the channel gating.

Early functional works showed that the *KCNK4* channel is involved in maintaining the resting membrane potential and is closely associated with neuron excitability^{50–52}. This gene is usually coexpressed with heat-sensitive and cold-sensitive excitatory channels, such as *TRPV1*, *TRPM8*, and *TRPA1*, and functions as a silencer of these excitatory channels in sensory neurons^{53–56}. It has been reported that downregulated expression or knockout of *KCNK4* in mice can increase neuronal sensitivity towards a given thermal stimulus, causing heat hyperalgesia or cold hypersensitivity in mice^{53,54}. This is because the deletion of neuronal K^+ channels in nociceptors will slow the depolarization of sensory neurons and prolong action potentials, thus amplifying painful sensations⁵³. Our protein structure simulation analysis suggested that the mutations in pit viper *KCNK4* channel are likely to affect channel gating, which might blunt the channel responses to certain types of gating commands or alter the thermal or mechanical sensitivity of the channel. If these mutations make the thermal activation threshold of pit viper *KCNK4* increase, this channel will be less active at lower temperatures at which *TRPA1* is activated. This functional effect is similar to knockdown or knockout *KCNK4* in pit vipers' infrared sensory neurons. If our functional speculation is correct, the pit vipers seem to use a mechanism similar to that underlies heat hyperalgesia in a mouse model generated by the deletion of *KCNK4* to increase their thermosensitivity. Of course, this hypothesized mechanism needs further confirmation with functional studies.

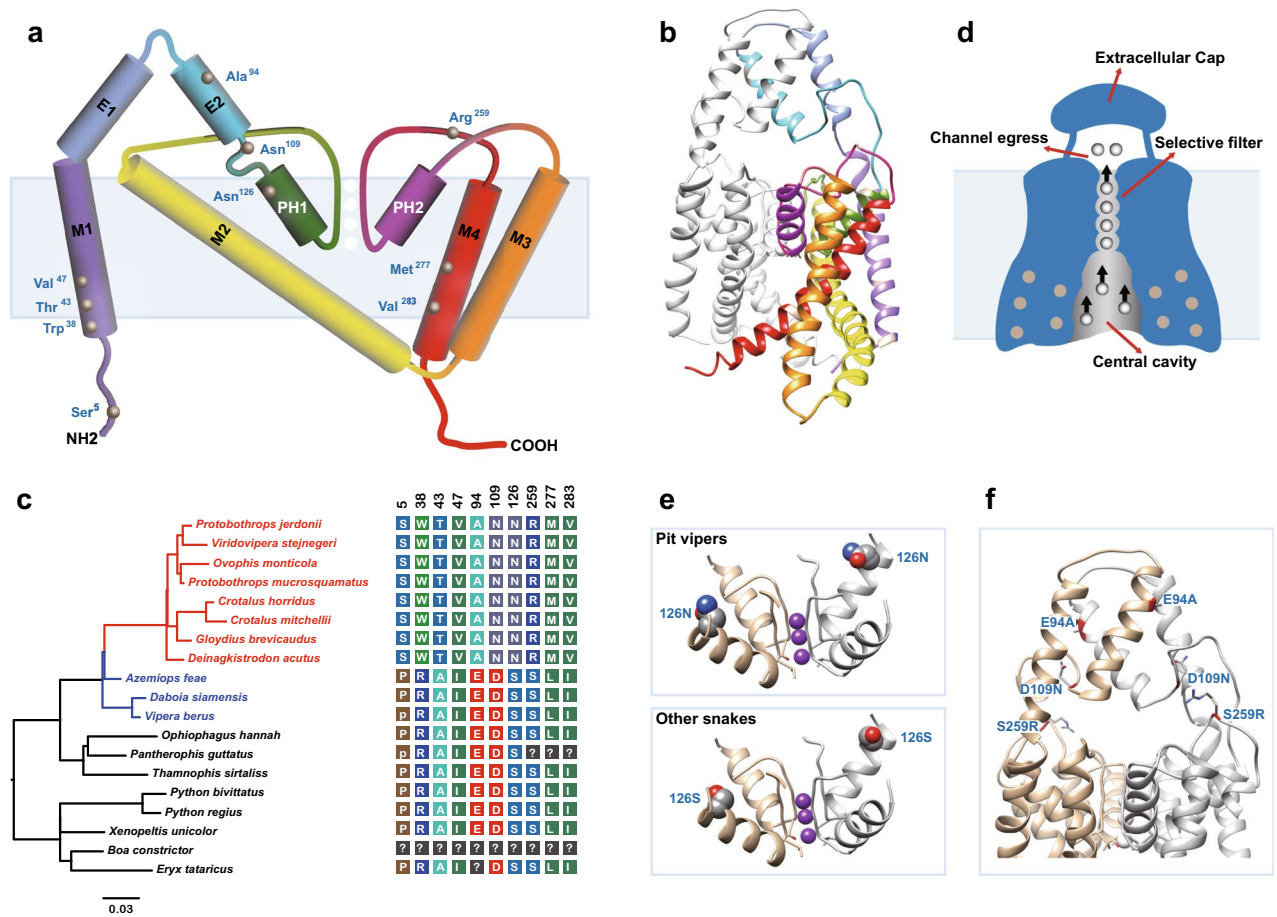


Figure 6. Protein structure simulation of the mutated pit viper KCN4 channel. **(a)** Topology diagram of KCN4 and the locations of the ten pit viper-specific amino acid substitutions. **(b)** Three-dimensional structure of the pit viper KCN4 dimer predicted by homology modeling based on the human KCN4 crystal structure. **(c)** The KCN4 gene tree of 19 snake species (pit vipers are red, and other vipers are blue) and a protein sequence alignment showing the ten pit viper-specific mutations (the full alignment is given in Supplementary Fig. S2). **(d)** Functional structure diagram of the KCN4 channel. Five amino acid mutations (R38W, A43T, I47V, L277M, I283V; labeled with filled grey circles) are present in the central cavity. **(e)** Structure of the selective filter in the two conformational states of the proteins from pit vipers (126 N) and other snakes (126S). **(f)** Three amino acid mutations (E94A, D109N, S259R) distributed around the KCN4 channel egress.

Conclusions

Compared with other specialized sensory systems such as mammalian echolocation and primate trichromatic vision, the molecular underpinning of snake infrared perception has received relatively little attention. Before this study, researchers had only performed transcriptome differential expression analyses to identify infrared perception-related genes in pit vipers and obtained only one target gene, *TRPA1*. On the basis of whole-exome capture sequencing, we conducted the first genome-wide mining of candidate genes associated with infrared perception function in pit vipers through comparative evolutionary analyses. Our study identified a total of 47 genes that underwent accelerated evolution, specifically in the pit viper lineage, which exhibited extensive accumulated pit viper-specific amino acid changes in their protein sequences. Most of the candidate genes were enriched in GO terms functionally related to ion transmembrane transport, which is closely related to infrared perception. We also examined the expression levels of these candidate genes in four different types of nerve tissues of both pit vipers and infrared-insensitive snakes and found that eight genes exhibited upregulated expression in the pit viper trigeminal ganglion. Furthermore, we simulated the protein structure of one of our candidate genes, the potassium channel gene *KCNK4*, and inferred its possible functional mechanisms. This work represents the first attempt to systematically identify genes associated with pit viper infrared perception based on comparative genomic analysis. We hope that the gene sets reported in this study provide a valuable stepping stone for further functional research and shed new light on the genetic basis of infrared perception in pit vipers.

Methods

Taxon sampling. In this study, a total of 19 snake species were sampled for comparative genomic analysis, covering all 3 subfamilies of Viperidae and representatives from five other relevant snake families (Supplementary Table S2). Of these snake species, nine have genomic data available in public databases, including four infrared-

sensitive crotalines (*Crotalus mitchellii*, *Crotalus horridus*, *Protobothrops mucrosquamatus*, and *Deinagkistrodon acutus*), one infrared-insensitive viper (*Vipera berus*), two infrared-insensitive colubrids (*Pantherophis guttatus* and *Thamnophis sirtalis*), one infrared-sensitive python (*Python bivittatus*) and one infrared-sensitive boa (*Boa constrictor*). In addition, we sampled another 10 snake species to conduct whole-coding-region sequence capture. These species included four Crotalinae species (*Gloydus brevicaudus*, *Protobothrops jerdonii*, *Viridovipera stejnegeri*, and *Ovophis monticola*) and one Azemiopinae species (*Azemiops feae*), one Viperinae species (*Daboia siamensis*), one Elapidae species (*Ophiophagus hannah*), one Pythonidae species (*Python regius*), one Xenopeltidae species (*Xenopeltis unicolor*) and one Boidae species (*Eryx tataricus*). This additional sampling supplements the missing Azemiopinae and enriches the sampling density of other families, especially pit vipers (Crotalinae).

All animal experiments in this study (including the RNA-seq experiments mentioned below) were performed in strict accordance with the guidelines and regulations developed by the China Council on Animal Care and Use. Animal processing procedures were approved by the Institutional Animal Care and Use Committee of Sun Yat-Sen University (permit number: 2018-023).

Whole-exome capture and sequencing. Genomic DNA of the 10 snake species were extracted from muscle tissues using the DNeasy Blood & Tissue Extraction Kit (Qiagen). The extracted gDNA samples were randomly fragmented by NEBNext dsDNA fragmentase (New England Biolabs) to 200–400 bp. A total of 500 ng of fragmented gDNA was used for DNA library construction with the NEBNext Illumina DNA Library Prep Kit (New England Biolabs) according to the manufacturer's protocol. We used the Custom Agilent SureSelectXT Target Enrichment System (Agilent Technologies) to capture whole coding regions from these libraries. The custom probes were designed at twofold tiling to target 21,331 snake CDSs for a total size of 32.1 Mb, based on the predicted exome of *Deinagkistrodon acutus*³². The hybridization enrichment experiments were performed using the SureSelectXT Reagent Kit (G9611A) according to the manufacturer's protocol. After enrichment, hybridized libraries were captured by Dynabeads MyOne Streptavidin C1 magnetic beads (Invitrogen) and amplified. The resulting captured libraries were pooled in equal quantities and sequenced on one Illumina HiSeq X-10 lane in 150-bp paired-end mode.

Orthologous gene group (OGG) construction. Raw sequence reads were quality filtered, adapters were trimmed using Trimmomatic v0.33⁵⁷, and the reads were then sorted to DNA samples based on library indexes. The filtered reads of each sample were aligned to the 21,331 *Deinagkistrodon acutus* reference CDSs using the MEM algorithm of Burrows-Wheeler Aligner version 0.7.10⁵⁸ with the default setting (-B 4) for viper species and relaxed setting (-B 2) for other snake species. The read mapping rate and the sequencing depth for each sample were calculated with SAMtools v1.4⁵⁹. Based on the BWA mapping results, the consensus sequences for each sample were assembled and called using SAMtools (mpileup command with -q 5 -Q 10 -B -uf) and BCFtools v1.4 (call command with -c). Using this bioinformatic pipeline, we obtained the orthologous CDSs of each sample corresponding to the 21,331 *Deinagkistrodon acutus* reference CDSs. For each reference CDS of *Deinagkistrodon acutus*, the obtained orthologous CDSs from different samples had exactly the same length, and the coordinates of these CDSs matched exactly those of the reference sequence.

Because some snake genome data used in this study were not well annotated, we used a so-called “pseudo sequencing strategy” to identify orthologous CDSs from snake species with published genome data. Briefly, the raw genome sequences of each snake species were artificially fragmented into small ‘reads’ of 150 bp at 10× tiling density. These “pseudo” reads were mapped to the *Deinagkistrodon acutus* reference CDSs and used to call orthologous CDSs using the aforementioned bioinformatic pipeline.

Finally, all the orthologous sequences for each reference CDS were put into a separate FASTA file as an orthologous gene group (OGG). The obtained OGG files were aligned using the ClustalW algorithm implemented in MEGA v6⁶⁰ according to the translated protein sequences. The program PAL2NAL v14⁶¹ was used to convert the protein sequence alignments into corresponding codon alignments for each OGG.

Phylogenetic analysis. To construct a genome-scale phylogeny for the 19 snake species sampled in this study, the protein alignments of all 21,331 OGGs were concatenated. The supermatrix was 11,231,292 amino acids in length with a data completeness of 90.3%. The phylogeny was reconstructed with maximum likelihood (ML) inference using RAXML v8.0⁶² under the GAMMA + WAGF model. Branch support was estimated by a rapid bootstrap algorithm (-f a option) with 500 replicates.

Branch length analysis. ML trees were separately reconstructed for the 21,331 OGGs based on their protein alignments using RAXML v8.0 with the GAMMA + WAGF model. To make branch lengths comparable across different OGG trees, we enforced the genome tree (Fig. 1a) as a topology constraint during tree inference (-g option) and estimated branch lengths.

Our branch length analysis aimed to search for genes that evolved faster in pit vipers (Crotalinae) than in the comparative snake group, totally including three steps of comparison: the first step compares Crotalinae (pit vipers) and Colubridae + Elapidae (non IR-sensitive snakes); the second step compares Crotalinae (pit vipers) and Viperinae + Azemiopinae (non IR-sensitive viper snakes); and the third step compares the ancestor of pit vipers and extant Viperinae and Azemiopinae (Fig. 2). The average branch length calculation was performed using an in-house Python script and data plotting was performed using R. To avoid repetition, the logic of these comparisons and their purposes are not described here but are given in the corresponding Results section.

Selection pressure analysis. Selection pressure analysis was performed using the CODEML program of the PAML package³⁹. To reduce the impact of missing data, all the OGG codon alignments were first filtered

to exclude species with more than 20% missing data. Only OGG codon alignments with more than three pit vipers and three non-pit snake species were used in the selection pressure analysis. For each OGG codon alignment, a separate user-tree was generated by pruning the corresponding missing species from the genome tree and labeling the ancestral branch of the pit viper lineage as the foreground branch. The two-ratio branch model (one ω was calculated for the foreground branch and another ω for all background branches; parameter setting "model=2, NSsites=0, fix_omega=0") was compared to the one-ratio branch model (estimated only one ω for all branches; parameter setting "model=0, NSsites=0, fix_omega=0"). After filtering out extremely low dN and dS values, the statistical *p*-values were calculated by likelihood ratio test (LRT) with one degree of freedom and false discovery rates (FDR) were computed using the Benjamini–Hochberg procedure⁶³. This analysis searches for genes that experienced significant selection pressure alteration in the early evolution of pit vipers compared to other snakes.

For genes showing signs of selection pressure alteration, we used the branch-site model in PAML (parameter setting "model=2, NSsites=2") to identify positively selected sites in the pit viper lineage. The null model (fix_omega=1, omega=1; all codons evolved neutrally) was compared with the alternative model (fix_omega=0, omega=1.5; some codons on the foreground branch experienced positive selection). The probable positively selected sites were identified and scored under Bayes empirical Bayes (BEB) probability.

Gene ontology analysis. To provide an overview of the over-represented functional categories of identified genes, we performed GO enrichment analyses with DAVID 6.7 (<https://david.abcc.ncifcrf.gov/>) using lists of all candidate genes that we identified from both selection pressure analysis and branch length analysis. Enrichment was determined using the Fisher's exact test and all *P*-values were adjusted for multiple comparisons by the Bonferroni method. Enriched GO categories were called using a 0.05 significance threshold and a minimum of 2 mapping entries.

Gene expression analysis. We generated new RNA-seq data for two types of central nerve tissues (cerebrum and midbrain) for a pit viper *Gloydus brevicaudus* and another comparable no-pit snake *Lycodon rufozonatus*. Total RNA was extracted from 200 mg of fresh nervous tissues using the RNA prep Pure Tissue Kit (Tiangen, Beijing). mRNA isolation, library construction, sequencing, data cleaning and quality control were performed by BGI (Shenzhen, China). All libraries were sequenced on an Illumina HiSeq2000 sequencer. We obtained approximately 3.5 Gb of 90-bp paired-end read data ($Q \geq 20$) for each tissue sample. RNA-seq data of peripheral nerve tissues (TG and DRG) were downloaded from a previous study¹⁸ (accession no. PRJNA121945; rattlesnake *Crotalus atrox* and combination data of rat snake *Elaphe obsoleta lindheimeri* and western coachwhip *Masticophis flagellum testaceus*). De novo transcriptome assembly was performed using Trinity v2014-07-17⁶⁴ with default parameters. And then homologous sequences were identified through a bidirectional BLAST (*e*-value $\leq 10^{-10}$). Reads from individual tissue libraries were mapped to the corresponding transcriptome assembly by Bowtie2 v. 2.2.6⁶⁵ with default parameters. Read counting for each transcript was performed by RSEM v. 1.3.0⁶⁶. Raw read counts were then scaled and normalized per sample by transforming to FPKM values, and cross-sample normalization was performed using a trimmed mean of M values (TMM⁶⁷) for comparison between libraries and species. Scaling and normalization were performed using the Trinity scripts "align_and_estimate_abundance.pl" and "abundance_estimates_to_matrix.pl", respectively.

Structural modeling. The structural models of two snake KCNK4 channels (pit viper *Deinagkistrodon acutus* and true viper *Vipera berus*) were generated based on the structure of human KCNK4 (PDB: 3UM7), with which these proteins share approximately 62% sequence identity. The SWISS-MODEL server (<https://swissmodel.expasy.org/>) was used to perform homology modeling of the dimer complex. The N terminus (amino acids 1–33) and C terminus (amino acids 299–464) were removed because no reliable coordinates are available for these regions in human KCNK4. The resulting 3D structures were graphically represented using PyMOL (<https://www.pymol.org>) and Chimera⁶⁸.

Data availability

The sequence capture reads generated in this study were deposited in the NCBI Short Read Archive under the BioProject accession number PRJNA511926, and the RNA-seq reads were deposited under the BioProject accession number PRJNA512390. All sequence alignments have been deposited as part of a Figshare project and are available at the following URL: <https://figshare.com/s/3d20ad7188b9070e7a1f>.

Received: 29 March 2020; Accepted: 14 July 2020

Published online: 03 August 2020

References

1. Noble, G. K. & Schmidt, A. The structure and function of the facial and labial pits of snakes. *Proc. Am. Philos. Soc.* **77**, 263–288 (1937).
2. Newman, E. A. & Hartline, P. H. The infrared vision of snakes. *Sci. Am.* **246**, 98–107s (1982).
3. de Cock, B. T. Thermal sensitivity as a specialization for prey capture and feeding in snakes. *Am. Zool.* **23**, 363–375 (1983).
4. de Cock, B. T. Thresholds of infrared sensitive tectal neurons in *Python reticulatus*, *Boa constrictor* and *Agkistrodon rhodostoma*. *J. Comp. Physiol. A* **151**, 461–467 (1983).
5. Ebert, J. & Westhoff, G. Behavioural examination of the infrared sensitivity of rattlesnakes (*Crotalus atrox*). *J. Comp. Physiol. A* **192**, 941–947 (2006).
6. Ebert J. *Infrared sense in snakes: Behavioural and anatomical examinations (Crotalus atrox, Python regius, Corallus hortulanus)*. Dr rer. nat. thesis, Rheinische Friedrich Wilhelms Univ. Bonn (2007).

7. Bakken, G. S. & Krochmal, A. R. The imaging properties and sensitivity of the facial pit of pitvipers as determined by optical and heat-transfer analysis. *J. Exp. Biol.* **210**, 2801–2810 (2007).
8. Clarke, J. A., Chopko, J. T. & Mackessy, S. P. The effect of moonlight on activity patterns of adult and juvenile prairie rattlesnakes (*Crotalus viridis viridis*). *J. Herpetol.* **30**, 192–197 (1996).
9. Kardong, K. V. & Mackessy, S. P. The strike behavior of a congenitally blind rattlesnake. *J. Herpetol.* **25**, 208–211 (1991).
10. Goris, R. C. Infrared organs of snakes: An integral part of vision. *J. Herpetol.* **45**, 2–14 (2011).
11. Bullock, T. H. & Diecke, F. P. J. Properties of an infrared receptor. *J. Physiol.* **134**, 47–87 (1956).
12. Terashima, S., Goris, R. C. & Katsuki, Y. Structure of warm fiber terminals in the pit membrane of vipers. *J. Ultrastruct. Res.* **31**, 494–506 (1970).
13. Berson, D. M. & Hartline, P. H. A tecto-rotundo-telencephalic pathway in the rattlesnake: Evidence for a forebrain representation of the infrared sense. *J. Neurosci.* **8**, 1074–1088 (1988).
14. Terashima, S. & Liang, Y. F. Temperature neurons in the crotaline trigeminal ganglia. *J. Neurophysiol.* **66**, 623–634 (1991).
15. Kohl, T., Bothe, M. S., Luksch, H., Straka, H. & Westhoff, G. Organotopic organization of the primary infrared sensitive nucleus (LTTD) in the western diamondback rattlesnake (*Crotalus atrox*). *J. Comp. Neurol.* **522**, 3943–3959 (2014).
16. Hartline, P. H., Kass, L. & Loop, M. S. Merging of modalities in the optic tectum: Infrared and visual integration in rattlesnakes. *Science* **199**, 1225–1229 (1978).
17. Newman, E. A. & Hartline, P. H. Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science* **213**, 789–791 (1981).
18. Gracheva, E. O. *et al.* Molecular basis of infrared detection by snakes. *Nature* **464**, 1006–1011 (2010).
19. Pyron, R. A., Burbrink, F. T. & Wiens, J. J. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evol. Biol.* **13**, 1 (2013).
20. Alencar, L. R. V. *et al.* Molecular phylogenetics and evolution diversification in vipers: Phylogenetic relationships, time of divergence and shifts in speciation rates. *Mol. Phylogenet. Evol.* **105**, 50–62 (2016).
21. Safer, A. B. & Grace, M. S. Infrared imaging in vipers: Differential responses of crotaline and viperine snakes to paired thermal targets. *Behav. Brain Res.* **154**, 55–61 (2004).
22. Roelke, C. E. & Childress, M. J. Defensive and infrared reception responses of true vipers, pitvipers, Azemiops and colubrids. *J. Zool.* **273**, 421–425 (2007).
23. Greene, H. W. *Snakes: The Evolution of Mystery in Nature* (University of California Press, California, 1997).
24. Preuss, T. M. Human brain evolution: From gene discovery to phenotype discovery. *Proc. Natl. Acad. Sci. USA* **109**, 10709–10716 (2012).
25. Enard, W. *et al.* Molecular evolution of *foxp2*, a gene involved in speech and language. *Nature* **418**, 869–872 (2002).
26. Zhang, J., Webb, D. M. & Podlaha, O. Accelerated protein evolution and origins of human-specific features: *Foxp2* as an example. *Genetics* **162**, 1825–1835 (2002).
27. Clark, A. G. *et al.* Inferring non-neutral evolution from human-chimp-mouse orthologous gene trios. *Science* **302**, 1960–1963 (2003).
28. Li, G. *et al.* The hearing gene *Prestin* reunites echolocating bats. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13959–13964 (2008).
29. Li, Y., Liu, Z., Shi, P. & Zhang, J. The hearing gene *Prestin* unites echolocating bats and whales. *Curr. Biol.* **20**, R55–R56 (2010).
30. Liu, Y. *et al.* Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* **20**, R53–R54 (2010).
31. Geng, J., Liang, D., Jiang, K. & Zhang, P. Molecular evolution of the infrared sensory gene *TRPA1* in snakes and implications for functional studies. *PLoS ONE* **6**, e28644 (2011).
32. Yin, W. *et al.* Evolutionary trajectories of snake genes and genomes revealed by comparative analyses of five-pacer viper. *Nat. Commun.* **7**, 13107 (2016).
33. Vicoso, B., Emerson, J. J., Zektser, Y., Mahajan, S. & Bachtrog, D. Comparative sex chromosome genomics in snakes: Differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol.* **11**, e1001643 (2013).
34. Aird, S. D. *et al.* Population genomic analysis of a pitviper reveals microevolutionary forces underlying venom chemistry. *Genome Biol. Evol.* **9**, 2640–2649 (2017).
35. Gilbert, C. *et al.* Endogenous hepadnaviruses, bornaviruses and circoviruses in snakes. *Proc. Biol. Sci.* **281**, 20141122 (2014).
36. Castoe, T. A. *et al.* The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20645–20650 (2013).
37. Ullate-Agote, A., Milinkovitch, M. C. & Tzika, A. C. The genome sequence of the corn snake (*Pantherophis guttatus*), a valuable resource for EvoDevo studies in squamates. *Int. J. Dev. Biol.* **58**, 881–888 (2014).
38. Vonk, F. J. *et al.* The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20651–20656 (2013).
39. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
40. Julius, D. & Nathans, J. Signaling by sensory receptors. *Cold Spring Harb. Perspect. Biol.* **4**, a005991 (2012).
41. Zou, A. R. *et al.* Distribution and functional properties of human KCNH8 (Elk1) potassium channels. *Am. J. Physiol. Cell Physiol.* **285**(6), C1356–C1366 (2003).
42. Salmikangas, P. *et al.* Myotilin, the limb-girdle muscular dystrophy 1A (LGMD1A) protein, cross-links actin filaments and controls sarcomere assembly. *Hum. Mol. Genet.* **12**, 189–203 (2003).
43. Noel, J., Sandoz, G. & Lesage, F. Molecular regulations governing TREK and TRAAK channel functions. *Channels (Austin)* **5**, 402–409 (2011).
44. Brohawn, S. G., del Marmol, J. & MacKinnon, R. Crystal structure of the human K_{2p} TRAAK, a lipid- and mechano-sensitive K⁺ ion channel. *Science* **335**, 436–441 (2012).
45. Lolicato, M., Riegelhaupt, P. M., Arrigoni, C., Clark, K. A. & Minor, D. L. Jr. Transmembrane helix straightening and buckling underlies activation of mechanosensitive and thermosensitive K_{2p} channels. *Neuron* **84**, 1198–1212 (2014).
46. Bauer, C. K. *et al.* Mutations in KCNK4 that affect gating cause a recognizable neurodevelopmental syndrome. *Am. J. Hum. Genet.* **103**, 621–630 (2018).
47. Brohawn, S. G., Campbell, E. B. & MacKinnon, R. Domain-swapped chain connectivity and gated membrane access in a Fab-mediated crystal of the human TRAAK K⁺ channel. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2129–2134 (2013).
48. Bagriantsev, S. N., Clark, K. A. & Minor, D. L. Jr. Metabolic and thermal stimuli control K_{2p}2.1 (TREK-1) through modular sensory and gating domains. *EMBO J.* **31**, 3297–3308 (2012).
49. Chatelain, F. C. *et al.* TWIK1, a unique background channel with variable ion selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5499–5504 (2012).
50. Fink, M. *et al.* A neuronal two P domain K⁺ channel stimulated by arachidonic acid and polyunsaturated fatty acids. *EMBO J.* **17**, 3297–3308 (1998).
51. Enyedi, P. & Czirjak, G. Molecular background of leak K⁺ currents: Two-pore domain potassium channels. *Physiol. Rev.* **90**, 559–605 (2010).
52. Lotshaw, D. P. Biophysical, pharmacological, and functional characteristics of cloned and native mammalian two-pore domain K⁺ channels. *Cell Biochem. Biophys.* **47**, 209–256 (2007).
53. Noel, J. *et al.* The mechano-activated K⁺ channels TRAAK and TREK-1 control both warm and cold perception. *EMBO J.* **28**, 1308–1318 (2009).

54. Descoeur, J. *et al.* Oxaliplatin-induced cold hypersensitivity is due to remodelling of ion channel expression in nociceptors. *EMBO Mol. Med.* **3**, 266–278 (2011).
55. Alloui, A. *et al.* TREK-1, a K⁺ channel involved in polymodal pain perception. *EMBO J.* **25**, 2368–2376 (2006).
56. Babes, A. Ion channels involved in cold detection in mammals: TRP and non-TRP mechanisms. *Biophys. Rev.* **1**, 193–200 (2009).
57. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
58. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
59. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
61. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucl. Acids Res.* **34**, W609–W612 (2006).
62. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
63. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.* **57**, 289–300 (1995).
64. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
65. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
66. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
67. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
68. Petersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank all of our lab members for help in experiments and data analyses. This work was supported by the National Natural Science Foundation of China (Grants No. 31872205 and 31672266 to P. Zhang), and the National Youth Talent Support Program to P. Zhang (W02070133).

Author contributions

P.Z. and D.L. designed the project. N.T. performed the sequence capture and RNA-seq experiments. N.T. analyzed the data and wrote the paper with the help of P.Z. and D.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69843-w>.

Correspondence and requests for materials should be addressed to P.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020